Psychological Bulletin

HARRY HELSON, Editor
University of Texas

CONTENTS

Developmental Studies of Perception. . . . . JOACHIM F. WOHLWILL 249

A Note on Projection: . . . . . PHILIP H. CHASE 289

Imprinting: Empirical Basis and Theoretical Significance. . . . . HOWARD MOLTZ 291

Contradictory Conclusions from Two Speed of Performance Measures. . . . . EUGENE S. EDGINGTON 315

Significance Tests for Multiple Comparison of Proportions, Variances, and Other Statistics. . . . . THOMAS A. RYAN 318

Published Bimonthly by the
American Psychological Association

VOL. 57, NO. 4 JULY 1960

Consulting Editors

R. R. BLAKE
University of Texas

D. W. MACKINNON
University of California, Berkeley

W. R. GARNER
Johns Hopkins University

L. J. POSTMAN
University of California, Berkeley

J. P. GUILFORD
University of Southern California

S. B. SELLS
Texas Christian University

W. H. HOLTZMAN
University of Texas

W. A. WILSON, JR.
University of Colorado

The *Psychological Bulletin* contains evaluative reviews of research literature and articles on research methodology in psychology. This JOURNAL *does not publish reports of original research or original theoretical articles.*

Manuscripts should be sent to the Editor, Harry Helson, Department of Psychology, University of Texas, Austin 12, Texas.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the *Publication Manual of the American Psychological Association* (1957 Revision). Special attention should be given to the section on the preparation of the references (pp. 50–60), since this is a particular source of difficulty in long reviews of research literature. *All copy must be double spaced, including the references.* All manuscripts should be submitted *in duplicate*, one of which should be an original typed copy; author's name should appear only on title page. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail and to check carefully the typing of the final copy.

Reprints. Fifty free reprints are given to contributors of articles and notes.

ARTHUR C. HOFFMAN
Managing Editor

HELEN ORR
Promotion Manager

SARAH WOMMACK
Editorial Assistant

Communications—including subscriptions, orders of back issues, and changes of address—should be addressed to the American Psychological Association, 1333 Sixteenth Street N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Annual subscription: $8.00. (Foreign $8.50). Single copies, $1.50.

PUBLISHED BIMONTHLY BY:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Menasha, Wisconsin

and 1333 Sixteenth Street N.W., Washington 6, D.C.

Entered as second class matter at Washington, D.C., and Menasha, Wisconsin. Second-class postage paid at Washington, D.C., and at additional mailing offices. Printed in U.S.A.

Copyright, 1960, by The American Psychological Association, Inc.

VOL. 57, NO. 4 JULY 1960

Psychological Bulletin

DEVELOPMENTAL STUDIES OF PERCEPTION[1]

JOACHIM F. WOHLWILL
Clark University

Until fairly recently interest in the ontogenetic development of perception on the part of English-speaking psychologists has been sporadic at best. Work on this problem has been generally ignored in the main literature in the field of perception, as in that of child development, although Vernon's (1954) survey of perception and Werner's (1957) comparative psychology treatise represent notable exceptions in this regard.

There are, however, signs of a definite reawakening of interest in the problem of perceptual development of late, paralleling perhaps the similar revival in the area of perceptual learning (e.g., Gibson & Gibson, 1955). Thus an increasing number of individual studies, as well as several more extended programmatic research projects, such as those by Edgren (1953), Gollin (1956, 1959a, 1959b) and Wapner and Werner (1957), are being reported which explore a variety of perceptual problems at different developmental levels. Similarly, a recent textbook on child development (Baldwin, 1955) has turned to the facts of perceptual and cognitive development for its theoretical framework, in sharp contrast to the consistent neglect of this topic in prior textbooks in this field.

This would seem, then, to be an auspicious time for a review of the extensive literature on this topic, much of which, being of foreign origin, is generally unfamiliar to American psychologists.

The present review is limited to experimental studies involving comparisons among two or more age groups. In order further to delimit as sharply as possible the area covered by this review, studies dealing on the one hand with tasks of a primarily cognitive, conceptual or psychomotor nature, and on the other hand with responses known to be direct functions of receptor mechanisms (e.g., temperature sensitivity, dark adaptation, etc.) were excluded.

Following a brief consideration of methodological problems, studies will be reviewed under the following headings: sensory thresholds, illusions, orientation and localization, the constancies, depth, form, number, movement, time, and perceptual learning. A concluding section will be devoted to a review and discussion of the major developmental trends uncovered in this review.

METHODOLOGICAL CONSIDERATIONS

A variety of methodological problems faces the investigator in this particular field. Since some of them have rarely been spelled out and are

[1] The survey of the literature reported in this article was carried out in part while the author was a postdoctoral National Science Fellow at the University of Geneva, Switzerland. The critical comments of Seymour Wapner and Morton Wiener are gratefully acknowledged.

frequently ignored in the design and execution of experiments, they are discussed here briefly.[2]

First of all, several problems arise with regard to the *sampling* of *Ss* in cross-sectional comparisons among different age groups. Apart from the obvious requirement of comparability among the samples representing the various age levels, the choice of the particular age groups to be included in an investigation deserves more careful attention than it is generally given. It is patently not valid to extrapolate trends obtained from a few selected age groups to the whole course of development, since age trends in this area are frequently discontinuous, nonlinear and even **U** shaped. Ideally, then, the investigation should include enough age groups to allow a determination of the total developmental trend over the age span under consideration.

Secondly, several points may be brought out with respect to the *statistical analysis* of developmental data. The ordinary analysis of variance is altogether insensitive to the order among the groups which are compared; thus it is relatively lacking in power when applied to the means of a number of groups representing points on an ordered variable such as age. By the same token the finding of significant age group effects does not itself convey any information regarding the form or even the presence of a consistent age trend. A decided improvement in this situation is the technique of trend analysis, particularly as developed by Grant (1956) to handle nonlinear regressions. In applying this or similar parametric techniques, one needs of course to be alert to heterogeneity of variance in the data, since variability between *Ss* frequently decreases with age, accompanying an increase in accuracy (intrasubject variability). Under certain circumstances, furthermore, even an increase in variability with age may occur (e.g., Klimpfinger, 1933).

Finally, in the majority of the studies that rely on one or another of the psychophysical methods, there are various factors basic to the judgment being made which the developmental investigator needs especially to keep in mind, since the contribution of such factors may not be constant across different age levels; thus they may in themselves account for, or at least seriously affect, the age changes observed in a given situation. For instance, the methods of limits and adjustment typically give rise to starting-position effects which have been shown to change in amount and even direction with age (Piaget & Lambercier, 1951b; Wapner & Werner, 1957). Consequently, even if the results from increasing and decreasing series are averaged, the PSE's thus obtained will be subject to varying amounts of error at the different age levels. This consideration suggests that in developmental comparisons of threshold values in particular these methods are of doubtful usefulness.

A different constant error, common to all methods in which a standard is compared to a series of variables, is the so-called "error of the standard" (Piaget & Lambercier, 1943a): the stimulus serving as the standard tends to be overestimated. This effect appears with particular prominence in constancy experiments

[2] A more extended treatment of the subject of experimental methods in the study of perception in children, by Eleanor Gibson and Vivian Olum, will appear as a chapter in a forthcoming handbook on methods in child psychology. The writer is indebted to Eleanor Gibson for a prepublication draft of this chapter.

(Akishige, 1937; Piaget & Lambercier, 1943b; Piaget & Lambercier, 1956a), but has been noted also in various other situations (Piaget & Lambercier, 1943a, 1953, 1956b; Rivers, 1905). Furthermore, a study by Tampieri (1955) indicates that the direction of the error of the standard may change from over- to underestimation, depending on the absolute size of the standard, which suggests an interpretation in terms of adaptation-level theory (Helson, 1948). If, as argued by Piaget and Lambercier (1943a), this error decreases with age—and a certain amount of evidence supports this assertion—it is clearly imperative to control for this effect when making comparisons among age groups.

A further constant error may arise when the *S* has to match a standard to one of a set of variables presented simultaneously in an ordered series. According to Lambercier (1946a), a strong central-tendency effect enters into such judgments, systematically shifting the choice of the variable in the direction of the middle of the series. This effect—which may be related to the error of the standard, as interpreted by Tampieri (1955)—again appears to decrease with age.

In conclusion, it is likely that a very large proportion of the research to be reported here is open to criticism on one or another of the above-mentioned grounds. This applies in particular to the extended series of studies by Piaget and his collaborators, which bulks rather large in our review, and whose shortcomings should thus be noted in brief. First of all, most of these studies employ a fairly unorthodox psychophysical method devised by Piaget and Lambercier (1943a) for use with children, in which the experimenter attempts to "zero in" on the PSE by alternating between variables on either side of it. Although this "clinical concentric method" permits a relatively rapid determination of an *S*'s PSE, in actual application it appears to be far from systematic and relatively lacking in precision.

A second point concerning Piaget's work is the general neglect of matters of experimental design (e.g., order of presentation of different experimental conditions) and of tests of significance. Closer attention to these points might perhaps have checked Piaget's general tendency to overinterpret his results and to engage in lengthy ad-hoc explanations for findings of frequently dubious significance—a tendency not calculated to enhance the readability of his reports (for example, cf. Piaget and Lambercier, [1953]). Indeed, this weakness tends to detract from the appreciation of the very real contribution which Piaget's work represents, in its investigation of a large variety of significant problems and in the ingenuity and thoroughness displayed in the manipulation of pertinent stimulus variables.

If in examining these studies, then, as well as those from other sources which are frequently no less subject to methodological criticism, a rather heavy strain may seem to be placed on the reader's willingness to "assume an attitude towards the mere possible"—as Goldstein and Scheerer (1941, p. 4) put it—this appears nevertheless to be a price well worth paying, in view of the interest of the problems and the theoretical issues which they will be found to raise.

Review of Experimental Studies

Sensory Thresholds

The evidence regarding age changes in absolute and differential sensitivity for various physical di-

mensions will be reviewed rather briefly, relying in the main on Peters' (1927) review of the literature on this topic. (Except for some studies on absolute sensitivity in infancy, reviewed by Munn [1955, ch. 8], not much relevant work appears to have been carried out more recently.) Peters cites certain functions for which sensitivity decreases with age, such as the upper and lower absolute thresholds of pitch, as well as the two-point threshold of skin sensitivity. Most differential thresholds, on the other hand, steadily decrease with age, i.e., discrimination improves. Particularly marked changes in this direction have been found for pitch, weight and saturation, as well as for the recognition of hue (the threshold being measured here in terms of the degree of saturation required). Visual acuity likewise improves with age, although the two studies cited by Peters differ with respect to the amount of improvement found.

The question of the discrimination of visual extents deserves special attention in view of its immediate relevance to many of the problems to be discussed. During the period from about 12 months to 3 or 4 years, substantial changes in differential sensitivity for size, width and area have been shown by Welch (1939a, 1939b); at the same time these studies indicate that by dint of prolonged practice over periods of from 4 to 12 months in the second year of life, discrimination can be brought up to the level of considerably older control *Ss*. After the age of about four or five years, however, little further change appears to take place in the ability to discriminate size or length. Thus, Giering (1905) found no difference between 6- and 14-year-old children in their ability to reproduce a distance between two points (the performance of children below the age of six is inferior, but this is quite likely due to the cognitive difficulty of the task). Estes (1959) likewise reports that the accuracy of size judgments for triangles, circles, and squares does not change in the age range from kindergarten to college students. Volkelt (1926) also obtained very closely comparable thresholds in his three age groups (three- to six-year-olds, an unspecified school-age group, and adults) for comparisons of three-dimensional objects, i.e., spheres. For lines and circles the youngest group did yield substantially higher thresholds than the other two groups (at all age levels the width of the threshold was least for spheres and greatest for lines).

These results for visual size judgments thus generally go counter to the substantial developmental changes in differential sensitivity for the dimensions of pitch, weight, hue, and saturation noted previously. This discrepancy may reflect the greater complexity of the proximal stimulus in the case of the latter dimensions, and the opportunity thus created for progressively finer differentiation of the relevant information in the stimulus, as argued by Gibson and Gibson (1955) in their discussion of perceptual learning.

Illusions

Due in part to the persistent interest shown by Piaget and his collaborators in the investigation of a large variety of illusions, this topic accounts for a considerable and perhaps disproportionate share of the literature on the development of perception. As we shall see, however, much of this work touches on problems of more general significance to perceptual development. In order to facilitate comparisons among studies on the same phenomenon, the results

TABLE 1

SUMMARY OF DEVELOPMENTAL STUDIES OF SELECTED ILLUSIONS

Illusion and Reference	Age Groups Included	Age Trend Observed	Remarks
Delboeuf, (Fig. 1a)			
Giering, 1905	6 and 14	Illusion less at 14	
Piaget, Lambercier et al. (1942)	5 to 12 & adults	Decrease (contrast shown mainly at 5–7 years)	Errors studied as function of size of circles (see text).
Rüssel, 1934	4 to 6½	Small, irregular changes	
Parallel Lines (Figs. 1b, 1c)			
Giering, 1905	6 and 14	No difference	Only four values of variable used.
Piaget & v. Albertini, 1950	5 to 9 & adults	Decrease, *in general*	Errors studied as function of dimensions of figure (see text).
Titchener Circles (Fig. 1d)			
Rüssel, 1934	4 to 6½ & adults	Small decrease from 4 to 6½	
Wapner & Werner, 1957	6 to adults	Increase (irregular)	Variables displayed in ordered series on chart.
Size-Weight			
Ciampi (see Rey, 1930)	3–6; 7–15; adults; aments; dements	Increase in % of *Ss* showing illusion	No psychophysical measures. For percentage of *Ss*, aments = normals; dements lowest.
Ohwaki, 1953	3; 4; 5; aments	Increase (as in Ciampi)	Aments = normals of same M.A.
Rey, 1930	5 to adults; aments	Increase to 9, then decrease	Adults not "naive." Illusion lowest in aments.
Müller-Lyer			
Binet, 1895	7–12 and 10–14	Less in older group	
Noelting, in press	5 to 10 & adults	Decrease	Practice effects studied (see under perceptual learning)
Piaget & v. Albertini, 1950	5 to 9 & adults	Illusion smallest for adults	Errors studied as function of dimensions of arrows (see text)
Piaget, Maire & Privat, 1954	4 to 10 & adults	Consistent decrease	Normal illusion compared to that in Fig. 3 (see under form).
Pintner & Anderson, 1916	6 to adults; aments	Decrease	Strong anticipation effects. Adult aments = 11-year normals.
Van Biervliet, 1896	12–16 & adults	Less for adults	Figure presented vertically; angles of arrows varied.
Walters, 1942	6 to 19	Decrease to 10; rise 15–19	Illusion *increased* on retest of 6–9 group 3 years later.
Wapner & Werner, 1957	6 to 19	Decrease to 12; rise 15–19	
Horizontal-Vertical			
Fraisse & Vautrey, 1956	6; 9; adults	Varies with type of figure and exposure time (see text)	Some differences between adults of differing educational backgrounds.
Rivers, 1905	Children & adults	Less for adults	Data show overestimation of standard (whether V or H).
Walters, 1942	6 to 19	Decrease	Trend confirmed in retest of 6–9 group 3 years later.
Winch, 1907	8 to 15; adults	Slight decrease	Based on drawings (V to be equaled to H).
Würsten, 1947	5 to 13; adults	Rise to 10, then decrease	Various figures used; lines mostly nonintersecting.

for the more frequently studied illusions are summarized in Table 1; some of these, involving assimilation and contrast phenomena, are reproduced in Fig. 1.[3]

Spatial assimilation and contrast effects. The work on the Delboeuf and parallel-line illusions indicates that effects of assimilation to contextual stimuli decrease with age (cf. Table 1). This is shown clearly in the study by Piaget, Lambercier, Boesch and von Albertini (1942) on the Delboeuf illusion (Fig. 1a), in which the relative size of the circles was systematically varied, yielding an **S** shaped function, as illustrated in Fig. 2. (The curves in Fig. 2 actually represent an idealized composite of five separate sets of curves

[3] A study by Hartmann and Triche (1933) on age changes for a number of familiar optical illusions has been omitted from consideration in this review, since the authors limited themselves to the determination of the *proportion of subjects* susceptible to the various illusions. Given the compelling nature of most of these, the almost wholly negative results obtained would seem to be of questionable value.

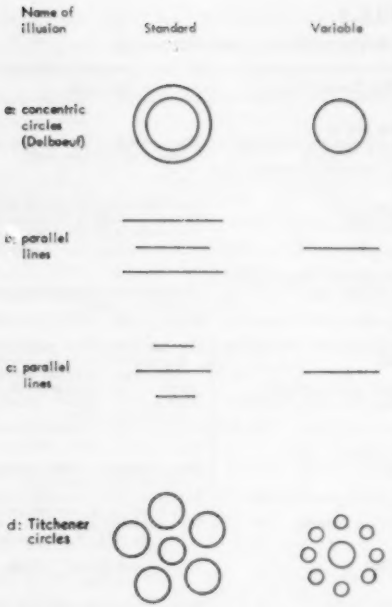


FIG. 1. Illusions of assimilation and contrast, as investigated in developmental studies cited in the text.

obtained for different absolute values of the stimuli. In particular, in three of the five cases the separate age curves differ substantially in the points at which they cross the X axis.) It will be noted that beyond a certain point further increase in the size of the context circle results in a contrast phase; however, only for the younger children does this phase appear consistently and prominently.

Very similar results were found by Piaget and von Albertini (1950) for the parallel-line illusion (Figs. 1b and 1c). Again, as the relative lengths of the inner and outer lines were varied, assimilation first increased, then decreased, changing eventually to contrast, though mainly for the children. Furthermore, variation in the vertical distance between the middle and the outer lines resulted in a change from a slightly negative to a strongly positive error (i.e., overestimation of the middle line for *both* figures). This latter result, indicating that this variable operated at the same time to increase assimilation (in Fig. 1b) and contrast (in Fig. 1c), together with the finding, corroborated in the Delboeuf study, that both assimilation and contrast effects decrease with age, suggest the possibility that both of these effects may be reducible to a single mechanism.

Piaget, in fact, has proposed such a mechanism, which he terms "centration," and which he has applied to a variety of illusions and related phenomena in perceptual judgment. This mechanism produces systematic distortions in perception, due to overestimation of the fixated elements of the stimulus field. In the course of development, however, these distortions are increasingly offset by "decentrations," resulting from more mobile and effective exploration of the total figure (cf. Piaget, Lambercier, Boesch & von Albertini, 1942).

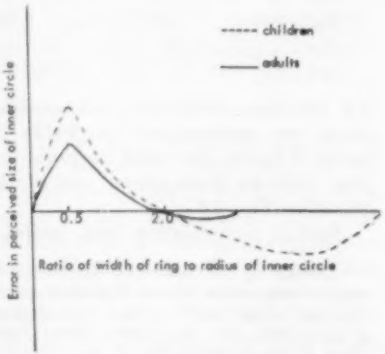


FIG. 2. Strength of concentric-circle illusion as a function of the relative width of the ring (after Piaget, Lambercier, Boesch & von Albertini, 1942).

Through further elaboration of these constructs Piaget has indeed arrived at a probabilistic model for perceptual judgment (Piaget, 1955) which he has successfully applied to a variety of illusions, all of which feature error curves of the type shown in Fig. 2 as a function of specific stimulus dimensions (e.g., Piaget & Pène, 1955; Piaget & Vurpillot, 1956). For our purposes, however, the main point is that according to Piaget's theory the shape of the error function remains constant with age, the extent of the illusion decreasing over all portions of the function. Thus both assimilation and contrast should decrease with age. Yet Wapner and Werner (1957), in their study on the Titchener circles (Fig. 1d)—an illusion at least superficially similar to those which we have been considering—found a contrast effect which increased with age. To what extent this discrepancy may be attributable to possible central-tendency effects arising from the exposure of the variables in an ordered series (cf. methodology section) and to what extent it reflects a more basic difference in the processes operating in this illusion is difficult to decide at this point. Systematic variation of the stimulus dimensions, as in the above-mentioned studies by Piaget and his collaborators, should lead to clarification of this question.

Temporal contrast effects. A study by Ikeda and Obonai (described in Sagara & Oyama, 1957) conclusively demonstrates that the Delboeuf illusion (Fig. 1a) can be transformed from an assimilation to a contrast phenomenon by changing the temporal relationship between the exposures of the outer and inner circles from simultaneity to succession. Developmentally there apppears to be a similar opposition between the simultaneous and successive conditions. This is indicated in Piaget and Lambercier's (1944) study of the Usnadze illusion, which is essentially equivalent to the Delboeuf illusion under successive presentation: two circles differing substantially in size are repeatedly exposed in a tachistoscope, following which a standard is shown within the area formerly occupied by the larger circle. A contrast effect develops (i.e., the standard is underestimated) which is greater in adults than in five- to seven-year-old children, in the sense that the illusion grows at a faster rate for adults as the number of presentations of the inducing stimulus increases. At the same time, however, the effect dissipates more rapidly for the adults after the final exposure of the inducing stimulus. It is of interest to note the correspondence of these findings with those of a recent study of figural aftereffects in normal and mentally deficient adolescents by Spitz and Blackman (1959), who found that their normals satiated more quickly, but that the effects also dissipated faster, in comparison with the mentally deficient group.

Piaget and Lambercier's explanation of the increase in contrast with age, in terms of adults' greater susceptibility to anticipation or set, receives some support from the similar developmental trend found for the *size-weight* illusion, which has usually been attributed to the contrast between the perceived weight of an object and that anticipated in view of its size. As Table 1 shows, susceptibility to this illusion increases markedly with age; it is also smaller in mental defectives as compared to normals of the same chronological age.

The Müller-Lyer illusion. This well-known illusion has proved very popular in developmental investigations, as seen in Table 1, which

shows, moreover, substantial agreement among the several studies as regards the decrease of this illusion with age. This decrease has generally been interpreted in terms of the younger child's difficulty in isolating the parts from the whole—a problem which will be more thoroughly discussed in the section on form perception. One empirical question which remains to be settled concerns the pattern of changes in this illusion through the period of adolescence. The studies by Walters (1942) and by Wapner and Werner (1957) both indicate a rise in the strength of the effect in later adolescence and early adulthood, whereas in the studies from the Geneva laboratories (Noelting, in press; Piaget, Maire, & Privat, 1954; Piaget & von Albertini, 1950) the adult groups yield the smallest average illusion. (None of the Geneva studies, however, included *S*s in the adolescent range.)

The investigation by Piaget and von Albertini (1950) is of special interest for the comparison made between the Müller-Lyer and the parallel-line illusion. Pointing to the formal similarity between the two types of figures (the latter may be derived from the former by eliminating the arrows and connecting their end-points), the authors have studied the changes in the Müller-Lyer as a function of the same stimulus variables manipulated in the parallel-line figures. These changes, as well as those due to age, follow a pattern which is closely similar to that found for the corresponding changes in the parallel-line illusion, even though in absolute terms the effects for the Müller-Lyer are much stronger. Thus the inference of common processes operating in the two situations appears justified.

The horizontal-vertical illusion. Despite a considerable amount of developmental work on this illusion (cf. Table 1), the nature of its change with age remains uncertain. In view of the differences among the various studies in the types of stimulus figures used, as well as in the methodology employed, and in view also of the failure to control systematically contaminating factors, such as the role of the bisection of the horizontal in the inverted **T**, the true meaning of the conflicting results obtained is difficult to assess.

The findings from the two major studies, by Walters (1942) and by Würsten (1947), nevertheless suggest the following tentative conclusion: if the two lines of the figure intersect (as they presumably did in Walters' study), the effect decreases with age; if, on the other hand, the *S* has to compare noncontiguous segments, as in almost all of Würsten's stimuli, the illusion increases up to a maximum at about age 10, subsequently declining up to adulthood—a trend which emerged very regularly and consistently from Würsten's data. In partial support of this statement one may cite the study by Fraisse and Vautrey (1956), who found an increase between 6 and 10 years for a figure made up of two spatially separated segments, but found no appreciable age trend for the inverted-**T** figure. (The two figures also differed markedly in terms of the effects of short vs. long exposure times.)

What might account for this discrepancy between these two types of figures? A clue may perhaps be found in Piaget's interpretation of Würsten's findings as reflecting the gradual development of a stable system of spatial coordinates in the age range from 6 to 10, resulting in an enhancement of the illusion (cf. Piaget &

Morf, 1956). But this factor would probably operate only if the two lines to be compared are not spatially contiguous, so that their spatial relationship is not clearly defined for the *S*. If the lines intersect, on the other hand, they provide in effect their own spatial framework, allowing for a more immediate operation of the illusion. The fact that the illusion is considerably stronger in the latter case (Fraisse & Vautrey, 1956) would support this argument.

Other illusions of orientation. If, as we shall demonstrate in the section on form perception, young children are relatively insensitive to the orientation of stimuli in space, one might expect them to be less susceptible than older *S*s to a variety of other illusions which depend on this factor. A case in point is the Schumann-square illusion, involving an overestimation of a square tilted at a 45-degree angle (i.e., a diamond), in comparison with a square of equal size oriented along the spatial horizontal and vertical. Hanfmann (1933) did in fact find an increase between the ages of three and six years in the percentage of *S*s responding to this illusion (after this age there was a decrease, attributed by the author to negative suggestion on the part of the older children). An attempt to measure the strength of this illusion by a psychophysical method did not result in any age differences, but the series of variables was apparently too coarse to yield very precise data.

In conclusion, several related studies by Piaget and his collaborators may be mentioned, which generally show an increase with age. Thus, in the comparison of two horizontal lines, staggered so as to form two sides of a parallelogram, the top line is overestimated increasingly between 5 and 8 years (Piaget & Taponier, 1956),[4] while Piaget and Morf (1956) find a similar age trend for the comparison between two similarly staggered vertical lines, as well as in the comparison of two oblique segments lying on a straight line. These various age trends are interpreted in terms of the increasing role of the spatial framework in the perception of the relationship between the lines. Accordingly, this age trend should not apply when two vertical lines are directly superimposed, since here the framework could not play an important role; indeed, Piaget and Morf (1956) find that in this situation the overestimation of the top line decreases with age. As the separation between the lines is increased, however, the overestimation becomes attenuated and the age differences correspondingly less consistent (Piaget & Lambercier, 1956b).

Miscellaneous illusions. A few illusions which have received only incidental attention will be mentioned briefly. Decreases with age have been found for the Sander parallelo-

[4] This study represents one of the very few instances (cf. also Piaget & Lambercier, 1946) in which Piaget has concerned himself with the relationship between perception and thinking in the context of an experimental study. In the present article he points to the contrast between the increase in error with age in the perceptual judgments and the marked decrease found over the same age range in the incidence of a purely cognitive error—lack of conservation of length—arising under identical stimulus conditions, when the actual equality of the lengths to be compared has been demonstrated to *S* beforehand.

It might be pointed out that Piaget appears in general to be more impressed by the differences than by the similarities between these two types of functions and their development; accordingly his work on perception is almost totally divorced from his more widely known work on conceptual development. He has, however, outlined the interrelationship between the two in more formal, theoretical terms (Piaget, 1950; Piaget & Morf, 1958).

gram (Heiss, 1930), the Poggendorff illusion (Vurpillot, 1957), the Oppel-Kundt segmented-line illusion (Piaget & Osterrieth, 1953), as well as for constant errors in the judgment of angles and of the lengths of their sides (Piaget, 1949; Piaget & Pène, 1955), in the judgment of chords of a circle (Piaget & Vurpillot, 1956) and in the comparison of the widths of rectangles varying in length (Piaget & Denis-Prinzhorn, 1953; cf. also Seashore & Williams, 1900). Increases with age appear in the case of the Jastrow illusion, which Giering (1905) found to be stronger at 14 than at 6 years, and for the Ponzo illusion (overestimation of a line which intersects a bundle of lines converging on a near-by point), for which Leibowitz and Heisel (1958) report an increase between four and seven years, though no further changes beyond that age. Finally, for an illusion of inclination, involving two staggered rows of rectangles separated by a horizontal line, which is perceived as tilted, the results of Piaget and Denis-Prinzhorn (1954) indicate in general an increase up to age 10, followed by a decrease, possibly reflecting processes similar to those operating in Würsten's (1947) study of the horizontal-vertical illusion. This effect is, however, strongly influenced by a variety of factors, such as the presence of a frame around the figure, the use of outline or solid rectangles, etc.

Before closing this section on the illusions, we should note a variable of particular interest manipulated by Vurpillot (1957) in her study of the Poggendorff illusion, that of the meaningfulness of the figure. When this author altered the figure so as to suggest a concrete scene, while leaving the basic stimulus elements unchanged, a very significant decrease in the illusion resulted for the 9-year, 12-year, and adult groups, whereas the two youngest groups (5 and 7 years) failed to profit from this alteration. While it would be dangerous to extrapolate this finding to the role of meaningfulness in general, it would seem to weaken the oft-heard argument that the poorer performance of young children in perceptual tasks is due to the use of abstract, meaningless stimuli.

Orientation and Localization

Spatial orientation. An extensive investigation has been carried out by Witkin et al. (1954) on the interaction of visual and bodily cues in the perception of the vertical at different age levels. These authors distinguish between two types of tasks: "field-as-a-whole" tests (e.g., *S*, in a tilted position, has to adjust a tilted room to the vertical) and "part-of-a-field" tests (e.g., *S*, in a tilted room, has to adjust his own position to the vertical). The first type of situation yielded no consistent age trends; on the other hand, in the part-of-a-field tests the *S*s' dependence on the visual framework seemed to decrease substantially with age. Thus in the tilted-room-tilted-chair test the tilted condition of the room has progressively less effect on *S*'s perception of himself as vertical. In interpreting this result, however, it is important to bear in mind the fact that even in the absence of conflicting visual cues the adjustment of the body to the vertical becomes increasingly accurate with age (Liebert & Rudel, 1959). Witkin's group likewise found a steady decrease with age in the influence of a tilted luminous frame enclosing a luminous rod in an otherwise dark room on the perception of the rod as vertical. But in a partial replication of this experiment by

Edgren (1953) no such age differences were found; instead, at all age levels (between eight and 16) the judgments ran *counter* to the direction of the frame.

A somewhat different emphasis is encountered in the work of Wapner and Werner (1957) who concern themselves with the influence of directional sets, induced visually or kinesthetically, on *S*'s orientation in space. These authors distinguish between three types of effects. The first of these involves *extraneous* stimulation, and is exemplified in judgments of verticality with the body in a tilted condition. Here the adult *S* overcompensates for the bias due to the tilt, so that the vertical is displaced in the opposite direction. But this tendency appears to develop only relatively late: prior to adolescence children show only very slight effects due to body tilt; for the girls, furthermore, the age curve crosses over at the youngest age level, where the vertical is displaced in the direction of tilt. Interestingly enough, this same displacement reappears in old age (Comalli, Wapner, & Werner, 1959). Independent confirmation of the developmental shift towards increasing overcompensation for extraneous stimulation, in the age range from early childhood to young adulthood, has most recently come from a similar study by Liebert and Rudel (1959) on the auditory localization of a moving tone with the body in a tilted position.

A second effect, *static object* stimulation, represents a bias due to the displacement of the object of judgment itself. It is illustrated primarily in starting-position effects: with the rod originally inclined to the left, the adjustment to the vertical deviates to the left, and vice versa for the rod originally inclined to the right. This effect decreases fairly regularly with age up to the college level, and furthermore does not increase again in old age (Comalli, Wapner & Werner, 1959). A further experiment reported by Wapner and Werner (1957), involving the adjustment of a luminous square to the apparent median plane, confirms the decrease with age in these starting-position effects, as does Liebert and Rudel's (1959) study on auditory localization.

The third of the effects considered by Wapner and Werner, *dynamic object* stimulation, applies to directional sets induced through the presentation of objects or words denoting particular directions (e.g., hands pointing upwards and downwards). For these effects the age trends were rather less consistent. Two of the studies on this problem disclosed, however, a developmental shift independent of these directional sets in the location of the apparent horizon, which was set at progressively lower levels with increasing age.

Spatial and cutaneous localization. A clearly related problem is that of spatial localization, which has been investigated in several developmental studies, giving at least indirect support to the increased effectiveness of the spatial framework. In the visual domain, Rey (1955) reports a steady decrease in the deviation of the reproduced location of a point from the true one, following a 5-sec. fixation period. Furthermore, if an extraneous stimulus—a point, a straight line, or a circle—is drawn on the sheet on which the point is to be localized, it will, for the younger children, exert a strong "pull," i.e., their localizations will be systematically shifted towards such a stimulus. This effect decreases with age, and may even become negative in adults. It would be of interest to determine the role of such

stimuli when they can themselves serve as referents for the localization to be made (i.e., when they are shown in the field in which the point is originally exposed, as well as in the subsequent one).

Sanström and Lundberg (1956) have failed to confirm the improvement in spatial localization with age, but since the response they used involved motor coordination to a much greater extent (the point, first shown on the top of a table, had to be found by sticking a pin on the underside of the table), it is scarcely comparable with Rey's. The role of the motor aspect is in fact shown by the age changes in the *direction* of the errors made: the children tended to "overshoot the mark," their errors being predominantly in the direction opposite to the hand used for the response.

Several studies are available on the development of cutaneous localization, but the results obtained are not too consistent. Dunford (1930) studied children between 3 and 15 years of age; he found only irregular changes in accuracy up to age 9; however, the largest mean errors were made by his 11- and 15-year-olds. Renshaw and Wherry (1931) failed to find any consistent age differences throughout childhood and adolescence; only his adult sample was clearly differentiated from the others, their mean errors being highest. The authors compare these results with those obtained under visual direction (i.e., eyes open, rather than blindfolded), which show a steady decrease in error.

Renshaw sees these results, together with those from a pilot study (Renshaw, 1930) which also showed adults to be inferior to children in cutaneous localization, as pointing to a changeover from kinesthetic to visual dominance in the perception of the body. In further support of this argument he adduces findings from a set of 11 *blind Ss*—seven older children and four adults—showing an improvement with age in cutaneous localization (Renshaw, Wherry & Newlin, 1930). But the small number of *Ss* used in all of Renshaw's studies (though his data represent means of a large number of trials for each *S*), together with the uncertain effect of the change in sheer anatomical dimensions with age on these absolute errors suggest caution in the interpretation of his results.

The Perceptual Constancies

The empiricism-nativism controversy has provided the impetus for a lively interest in the development of the constancies, and in particular that of size constancy, as well as in the related topic of depth perception. Some of the major issues in this area, however, confront us equally in the rather more compact literature on brightness constancy, which thus may serve as a convenient starting point.

Brightness. The two most important studies on developmental changes in brightness constancy, by Brunswik (1929) and Burzlaff (1931), represent sharply divergent positions on the question of the evolution of the constancies. Brunswik, emphasizing the role of experience in bringing about the perceptual achievement of the organism reflected in constancy, finds a steady increase in constancy between ages 3 and 11, followed by a slight but consistent falling off after that point. This reduction in constancy between early adolescence and adulthood, which is confirmed by the data of Walker (1927), is interpreted by Brunswik (1956) in functional terms, as indicating a partial supplanting of percep-

tion by intellectual functions in the adult.

Under experimental conditions closely comparable to Brunswik's, Burzlaff (1931) did substantiate the increase in constancy with age, at least between four and seven years (it should be noted that this study is limited to 20 children in this age range, together with a sample of five adults). But when the variable, instead of being presented as an isolated color wheel, was displayed in an ordered series of greys on a chart, the age trend disappeared, and at the same time constancy improved generally. This held true whether a color wheel or a random series of greys were used for the standard. Furthermore, Akishige (1937) has shown that if color wheels are used for both standard and variable, merely making a chart of greys available does not alter the situation: the age trend reappears, and constancy generally deteriorates, in comparison with judgments in which the variable itself is part of an ordered series.

The discrepant results found by Burzlaff and Brunswik gave rise to some heated debate as to the relative validity of the two sets of findings. Brunswik (cf. note to Klimpfinger, 1933, pp. 619ff) argued that Burzlaff's situations with serial presentation of the variable were too simple to allow age differences to manifest themselves; Burzlaff (1931) maintained, on the contrary, that Brunswik's experimental conditions were too artificial and unstructured to be meaningful for the children. This discussion misses what seems to be the main point, namely, the demonstrated role of the perceptual situation, and more particularly of the difficulty of the judgment for the *S*, as a determinant of developmental differences.

Shape. Klimpfinger (1933), in a study carried out under Brunswik's auspices, found a developmental trend rather similar to that found by Brunswik for brightness constancy: an increase up to about 14 years, followed by a subsequent decline, more pronounced than that in the brightness study: the mean of the older of the two adult groups corresponds roughly to the 8-year-old level. There is, however, some suggestion of bimodality in the distribution of the values at the older age levels.

Akishige (1937) has verified the increase of shape constancy between ages 4 and 10. At the same time he finds that this age change disappears when the variable is shown in a series of ellipses of varying eccentricity, i.e., under conditions comparable to those of Burzlaff for brightness constancy. Constancy is again generally higher under these latter conditions.

Size. The developmental investigation of size constancy has been far more thorough than for brightness and shape, both in terms of the manipulation of relevant stimulus variables and in terms of the extension of the work to infants and preverbal children. As we shall see, however, the methodological problems in the way of adequate testing of constancy in infants are formidable, and the results obtained are accordingly of dubious value.

Thus, an oft-cited study by Cruikshank (1941) purports to show that constancy, though not present at birth, appears by about 6 months. This conclusion is based on the fact that after that age infants ceased to respond consistently to a rattle at 75 cm., while they did respond to one equal to it in *projective size*, at a distance of 25 cm. But clearly this comparison only indicates discrimination of depth, i.e., a tendency to

ignore objects that are relatively far away. Indeed, if this method allowed any inference concerning constancy, one would have to conclude that it was absent altogether, since responses to a rattle at 75 cm. equal in physical size to the one shown at 25 cm. were infrequent at any age level, and decreased steadily after 15 weeks.

Cruikshank's procedure appears inherently unsuited to answer the question of constancy, since only a single rattle was shown at any given time. More recently, Misumi (1951) has approached the problem through a stimulus preference method, with somewhat different results. The main feature of his method is his determination of the extent of the *S*'s spontaneous preference for one of two objects differing in size, when presented at the same distance from the *S*. This is then used as a baseline in the evaluation of the results for pairs of objects of which one is removed in depth. Thus, from the age of five months on, 77% of the responses are made to the larger of two goldfish; when it is moved back so as to be equal in retinal size with the smaller one, the choices of the larger drop to between 9% and 21%, in different age groups between 20 and 59 weeks. Thus constancy appears to be absent here, *if* we can assume that the infant is still responding to size rather than depth. In a second experiment, however, in which the stimuli were small red balls, Misumi did find fairly good evidence for constancy, at least after the age of nine months: the extent of preference for the larger ball was unaffected, after this age, by displacing it in depth.

More satisfactory and unambiguous results can be obtained from nonverbal *Ss* by first establishing a stable differential response to size and then testing its maintenance as the distance of the stimuli is manipulated. This approach has been applied (Beyrl, 1926; Frank, 1925), though not in very systematic fashion; Frank's study is nevertheless of interest in demonstrating that one- to two-year-old children can consistently adhere to their size discrimination even with the larger stimulus as far as 15 m. removed from the smaller one. Beyrl reports, however, that an attempted replication of Frank's work resulted in a considerably higher incidence of errors.

The major portion of Beyrl's (1926) investigation consists of a psychophysical study of size constancy at different age levels, which, in contrast to the work reported heretofore, yields quantitative measures of the degree of constancy attained. His *Ss* were 2½- to 11-year-old children and a group of adults; they were given cubes, as well as discs, to judge as to size by a method of constant stimuli, with the standard at 1 m. and the variables at distances from 2 m. to 11 m. The age curves for the judgments at the separate distances of the variable all show a steadily decreasing tendency to underestimate the far stimulus, becoming negligible by age 11; as distance increases, furthermore, the younger *Ss* lag further and further behind the older ones in their approach to constancy. This is true both for the cubes and for the discs, but to an apparently greater extent for the latter.

Although these results clearly show that size constancy is far from perfect at the younger ages, this study has been cited repeatedly as demonstrating nearly 100% constancy even at the lowest age level. The basis for this misstatement is the graph through which Brunswik has compared this study with others of color and shape constancy (see Klimp-

finger, 1933; also Brunswik, 1956, p. 83); and which indicates a size-constancy ratio as high as .92 at the age of two and one-half years. This figure is, however, in large part an artifact of the arithmetical properties of the Brunswik ratio, which is virtually insensitive to any except very drastic errors at longer distances, as reflected in the paradoxical finding that the values of this ratio actually increase with distance for the younger children! If we add to this observation the disquieting though generally ignored fact that the results from 19 of the 55 preschool children had to be discarded, as the judgments were too inconsistent to allow the calculation of a PSE, the departure from constancy in the judgments at these lower age levels is seen to be quite considerable.

While virtually all of the subsequent studies on size constancy have substantiated this increase in constancy with age, several variables appear to be important determiners of the extent and even the presence or absence of this age trend. Thus, just as in the case of shape and brightness constancy, the presentation of the variable in an ordered series tends to equalize the performance of the different age groups, and improve constancy generally. This was shown by Burzlaff (1931) in an exploratory study modeled after his brightness-constancy work discussed above. On the other hand, in a reversal of this situation—variable shown singly and standard in an ordered series—Akishige (1937) did find a marked age trend between 4 and 10 years, suggesting that Burzlaff's effect is more specifically connected with the judgmental process involved in the selection of the variable to be equated to the standard.

This is in fact the position taken by Lambercier (1946a), in a paper which undoubtedly presents the most exhaustive and searching analysis available of methodological factors in the size-constancy problem. He believes that Burzlaff's and Akishige's results are to be attributed to the operation of a central tendency effect, already referred to in the methodology section, which is especially marked in the children, systematically displacing their judgments towards the middle of the series of variables. Where, as in Burzlaff's case, the standard is equivalent to this middle, the net effect would be to counteract any constant errors due to imperfect constancy. In the experimental portion of his study, Lambercier presents data which seem to bear out his thesis: as the center of the series of variables is shifted away from the value of the standard, the errors are systematically displaced in the direction of this center, but to an extent decreasing with age. Unfortunately the more eccentric variable series led at times to refusals on the part of *S* to admit any stimulus as equivalent to the standard, with awkward consequences not only for the calculation of PSE's, but particularly for the validity of the age comparisons, since such refusals occurred chiefly in adults. Furthermore, Lambercier's hypothesis cannot account for Burzlaff's brightness constancy data, since there the standard was generally eccentric with respect to the series of variables. It may be noted that Lambercier finds consistent increases in constancy across a variety of other experimental conditions; again developmental changes were most pronounced under single presentation of standard and variable, although they were attenuated due to practice effects upon repetition of the single-variable trials following judgments

under other, apparently easier conditions.

A further study by Lambercier (1946b) provides information on the role of reference stimuli in the visual field. A number of sticks were placed horizontally or vertically along the line of sight between variable and standard. The only condition materially affecting the size judgments was one in which the vertically placed reference sticks were all of equal size, and equal to the standard as well, thus enhancing linear perspective cues. The beneficial effect of this condition increased with age; however, since the equivalence of the reference sticks to the standard was demonstrated to the *S*s, this may actually represent more a cognitive than a strictly perceptual performance; the fact that 94% of the adult judgments were exactly correct suggests as much.

An additional factor relating to the stimulus conditions which has received some attention in size constancy research is that of the horizontal separation between the stimuli. Frank (1927), commenting on Beyrl's (1926) attempted replication of her earlier study (Frank, 1925), attributed the relatively poor performance of Beyrl's *S*s to the smaller horizontal separation between the two stimuli, which she believed would work against the perception of their objective size in young children. Comparing additional results obtained both under Beyrl's conditions of horizontal separation and under greater separation—the extent is not specified—she obtained in fact superior results for the three- and four-year-old *S*s under the latter conditions, though little difference for older children.

Piaget and Lambercier (1943b) have investigated this problem more systematically; they find an over-all increase of the average error under wide lateral separation between standard and variable, but little age difference in this respect. But the lateral separation used was rather extreme; in fact, depth and lateral separation were confounded here, as the variable was presented 3 m. to the *side* of the standard, with *S* directly behind one of the two.

A much more important factor appears to be the position of the standard and variable relative to the *S*. Both Akishige (1937) and Piaget and Lambercier (1943b, 1951a, 1956a) find higher degrees of constancy when the standard is in the far position. In fact, Piaget and Lambercier find overconstancy at all age levels in this situation, while this appears only after the age of 8 years when the standard is near (cf. Piaget & Lambercier, 1956a, Table 4, p. 277). This effect represents of course a special case of the "error of the standard" discussed in the methodology section. (It may be noted that both Burzlaff [1931] and Akishige [1937] find a similar effect in brightness constancy: when the standard is in the dark, constancy is considerably enhanced.) The differential effect of this error at different age levels does not emerge very clearly from these studies, but it is obviously a factor to be taken into account in interpreting a given constant error in terms of its absolute value.

Finally, the variable of *distance* needs to be considered. First of all, the relative distance between the stimuli has recently been shown to be a significant determinant of developmental differences. Cohen, Hershkowitz, and Chodack (1958) found a significant increase in constancy between 5 and 17 years when a standard at 2 m. was compared to a variable

at 8 m.; when the former was moved back to 6 m. from *S*, the age differences disappeared. As might be expected, there is a general reduction of error in this situation, and correspondingly less room for developmental change. For both situations, however, these authors found a significant reduction with age in the interval of uncertainty (i.e., the width of the DL), which is confirmed in the studies of Lambercier (1946b) and Piaget and Lambercier (1943b), for similar conditions involving single presentation of the variable.

Perhaps a more interesting factor is the role of the absolute distance of the stimuli from *S*. Its investigation would throw light on the developmental aspects of the so-called secondary cues to distance emphasized by Gibson (1950). This question has received little study thus far, however. Zeigler and Leibowitz (1957), in an experiment modeled after the classical study of Holway and Boring (1941), did determine size constancy at distances up to 100 feet, in an indoor setting. The results, from only 13 *S*s (eight boys aged seven to nine and five adults) showed that the judgments of the boys conformed to the law of the retinal angle to a surprising extent, whereas the adults, in agreement with Holway and Boring's findings for full-cue-conditions, show overconstancy (at least up to 60 feet).

Developmental changes in size constancy in a Gibsonian, full-texture field are the subject of one of the experiments of the thesis by Edgren (1953). His stimulus materials consisted of a series of photographic slides on which a standard stick was displayed at a distance in an empty field, with a series of variables in the foreground (cf. Gibson, 1950, pp. 184f). Distance and size of the standard varied from slide to slide, but the results are given only for the combined series; they indicate only very irregular age changes between 8 and 16 years. If the central-tendency effect postulated by Lambercier (1946a) under these conditions of serial presentation of the variable is a real one, the size of the standard may have been an important factor.

One analysis which provided Edgren with significant age differences in this experiment concerned judgments on two especially prepared slides in which the context had been masked completely. The effect of the loss of cues proved relatively greater for adults, which is just the opposite of what one might expect. Again central tendency effects, as well as more general uncontrolled variables (all of Edgren's experiments were group-administered) play unknown roles here.

Although not varying distance systematically, a study of Dukes (1951) is of incidental interest here, since it represents an adaptation of Brunswik's (1944) "ecological representativeness" approach, in which *S* judged the size of a large number of objects in his natural environment, over a large range of sizes and distances. The results for a single six-year-old child actually yield a higher correlation between judged and objective size (.99) than that obtained by Brunswik for an adult for a comparable set of stimulus objects (.95). An important difference in the two studies, however, is Dukes' use of comparative judgments, as opposed to the absolute-judgment procedure of Brunswik.

Rounding out this review of the work on size constancy are a few studies employing instructions for projective- (retinal-) size matches. Piaget and Lambercier (1951a, 1956a) have made an extensive study of such

judgments in their most typical constancy situation (standard at 1 m., variable at 4 m. or vice versa). The rather surprising result was that the younger children, aged seven to eight, provided they could understand the instructions at all, performed considerably better than the 10- to 12-year group, and even somewhat better than the adults. Thus there was a **U** shaped trend here, as in Würsten's (1947) study on the horizontal-vertical illusion. The authors suggest that for the youngest children there is less interference from the perception of the true size of the distant object, due to their tendency to perceive objects independently of the spatial framework. With increasing age this framework becomes increasingly determining, so that projective size becomes progressively harder to abstract, until the intervention of cognitive mechanisms mediating the perception of perspective relationships become available (i.e., in the adult). Whatever the merits of this explanation, the superiority of the youngest group remains somewhat mysterious, particularly since they were at the same time the most accurate in their objective-size matches (the older *Ss* exhibited overconstancy). It should be pointed out, incidentally, that even the best judgments of projective size fell far short of the actual projective match; at all ages there was thus a strong effect due to objective stimulus size.

The problem of projective-size estimates has been attacked in a rather different fashion by Weber and Bicknell (1935), who had their *Ss* (8 aged 9 to 12 and 20 adults) make phenomenal-size matches of discs photographed at varying distances; the photographs were viewed both under ordinary and stereoscopic conditions. In general, the instructions were apparently effective in leading to projective-size judgments, as reflected in constancy ratios between .18 and .25. The children showed a greater influence of the actual, as opposed to the photographic size of the objects, in agreement with the findings of Piaget and Lambercier for a comparable age range. At the same time, as regards the change in judgment with increasing distance, the results for the children were between those for the adult men and women.

Distance-constancy. In his discussion of the size-constancy problem, Gibson (1950, pp. 163ff) points out that not only the size of the objects, but the distances between them tend to remain constant as the individual moves through his normal environment. Massucco Costa (1949) has explored this question developmentally, asking *S* to chose a variable interval at 4 m. equal to a standard interval marked off at 1 m. The intervals were either empty (delimited by cubes or rulers) or filled. A developmental trend was in evidence only for the empty-interval judgments, which fell considerably short of constancy for five- and seven-year-old children (though not so much for the six-year-old group), while older children and adults exhibited nearly complete constancy.

A similar study by Denis-Prinzhorn (1959) is of particular interest, since it focuses on the question of the correlation between size and distance perception. In her experiment —which forms part of a dissertation currently in progress at Geneva—*Ss* made judgments both of the relative size of two sticks at 1 m. and 4 m., and of relative distance over the same stimulus field (the latter by a bisection procedure). Denis-Prinzhorn finds a developmental trend

from under- to overconstancy for both the size and the distance judgments. More important is her demonstration that these two judgments, which are virtually uncorrelated for adults—thus confirming similar results obtained by Gruber (1954)—are on the contrary highly correlated for young children (five to seven years). This finding suggests that cues which in adults operate relatively independently in the two situations are much more closely interrelated in the child. Exactly which cues are involved in this developmental change remains to be determined, however.

Concluding remarks on the constancies. The general picture with respect to the development of the constancies may perhaps be summarized as follows: *some* "regression to the real object"—to use Thouless' (1931) expression—appears to develop very early in life, probably by the end of the first year, although most of the data on this question are unreliable. Under conditions in which constancy is enhanced generally (e.g., serial presentation of the variable; short distances between standard and variable), the progress of development is almost complete at a fairly early age, i.e., five or six years. Further increases in constancy occur throughout later childhood and adolescence in relatively impoverished laboratory situations and with single presentation of the variable; this trend typically culminates in overconstancy in the adult.

It appears, then, that the young child requires a greater variety of cues than the adult in order to maintain invariance in his perception. This conclusion is also in conformance with the reduction with age in the correlation between size and distance judgments (Denis-Prinzhorn, 1959) and with the apparently sharper rise in constancy up to early adolescence for shape and brightness constancy, where the available cues are not so manifold, as compared with size constancy (Klimpfinger, 1933). One among the many problems that remain to be clarified in this regard, however, is the meaning of the decrease in shape and brightness constancy after early adolescence.

Depth Perception

While depth perception is clearly implied in the above-mentioned study on distance judgment by Denis-Prinzhorn (1959), the discrimination of two stimuli as relatively nearer or farther represents a somewhat different problem, and one less obviously related to size constancy. Gross differentiation of depth appears to develop quite early, according to the recent work by Walk and Gibson (1959), who have demonstrated that 6- to 14-month-old infants will rather consistently avoid a "visual cliff" 3 m. deep, on the basis of perceptual cues to depth. As we saw in our discussion of size constancy, the results obtained by Cruikshank (1941) likewise suggest strongly that after the age of six months infants regularly respond to the nearer of two objects when their separation in depth is considerable.

As for more finely differentiated judgments of the relative distance of two objects, an extended investigation by Updegraff (1930) shows that this type of discrimination is rather well developed already in four-year-old children, who are only slightly inferior to adults. This applies both to short (5 m.) and long (195 m.) distances of the objects from the observer. Furthermore, neither the children nor the adults show any appreciable differences between monoc-

ular and binocular viewing. On the other hand, under the presumably more difficult conditions of a Bourdon-type judgment (the *Ss*, looking through a small aperture, judged whether a ball dropped in front or behind a partition), the children were consistently less accurate than the adults.

This last type of judgment is of course a function of retinal disparity. Developmental changes in the effectiveness of this cue are shown in several studies of stereoscopic perception (Carr, 1935; Gesell, Ilg, & Bullis, 1949; Leyer, 1939). Carr (1935) reports that below the age of three years children were unable to perceive depth in stereoscopic pictures, even after a period of training; Gesell, Ilg, and Bullis (1949) find similar difficulties up to the age of 5. But, as Munn (1955, p. 269) points out, these results are of equivocal significance, in the absence of adequate control over *S*'s ability to focus effectively stereoscopic images. That this factor may play an important part is indicated by the contrasting results obtained by Johnson and Beck (1941). These authors utilized stereoscopic slides projected as polarized images on a screen (i.e., as in "3-D" motion pictures), which the child viewed through polaroid glasses. If the projected image of the object were perceived in depth, it would be localized at a point between *S* and the screen. Although the criteria for *S*'s perception of depth are not made too explicit, the authors' findings seem to indicate that even the youngest *Ss* (aged two years) did perceive the image three-dimensionally.

While the foregoing studies involved images of actual three-dimensional objects, Leyer (1939) used as his stimulus materials a set of stereoscopic geometrical drawings, shown under tachistoscopic exposure. In an 8- to 13-year-old group only 40% of the drawings were perceived three-dimensionally, as against 60% in a 15- to 19-year-old group. These two groups likewise responded differently to a task in which a vertical line, viewed stereoscopically, was to be adjusted so as to be midway between two other parallel lines. If the lines were all perceived in the same plane, the adjustment would naturally differ from that made if the middle line were perceived as displaced in depth. Here again a somewhat larger proportion of the older *Ss* responded in terms of depth.

Lastly, we may consider a rather different aspect of depth perception, namely the three-dimensional interpretation of two-dimensional figures. Very little work has been carried out on this question developmentally. The Gestaltists' assumption of the intrinsic determination of such depth effects receives some indirect support, however, from studies by Slochower (1946) and Ishii (1956). The latter asked for descriptions of **Y** and **T** shaped drawings, by *Ss* 6 and 14 years of age; no differences were found in the responses given by the two groups. Slochower used a wider range of stimuli, though less systematically constructed, again finding little age differences in the range from five to nine years in *Ss*' reproductions of the designs by means of modelling clay. In both of these studies the stimulus characteristics appeared to be the primary determinants of the perception of depth.

Form Perception

Form discrimination in infancy. As in the case of the constancies, the question of the perception of form at very early age levels has attracted a fair amount of attention. To antici-

pate the main trend of the findings to be presented, it seems that, while form discrimination can be demonstrated fairly early in infancy, it is not a very potent aspect of the stimulus in the perception of young children.

Ling (1941), by utilizing a discrimination-learning procedure in which the stimuli (sugar-coated three-dimensional forms) coincided with the rewards, was able to show discrimination between various simple geometrical shapes as early as six months; the differential response was furthermore little affected by changes in size or orientation of the correct figure. Of particular interest is Ling's finding that an initial difficult discrimination between a circle and an ellipse led to greater transfer when the ellipse was replaced with angular objects than that found for the reverse order (circle vs. square, followed by circle vs. ellipse).

This study thus shows that the mechanisms for shape discrimination are functional at an early age, or at least can become so following exposure to a limited amount of specific training (it should be noted that in this experiment, which extended over a period of several months for any given infant, learning and maturational factors are to some extent confounded). A probably important contributing factor to such early discrimination is the use of three-dimensional objects, since *size* discrimination is typically better and develops earlier for three-dimensional as compared with equivalent plane figures (Stevenson & McBee, 1958; Volkelt, 1926; Welch, 1939b).

Rüssel (1931) compared a variety of cues (contour vs. solid form, size, thickness, round vs. pointed form and symmetry) with respect to speed of learning of a discrimination and the extent of transfer to other stimuli differing on the same cue. The number of *S*s (24, aged one and one-half to five years) was too small to allow systematic age-group comparisons, but it is of interest that, although the cues were all roughly equivalent as regards the ease of the original discrimination, the transfer was considerably less for the last three cues—representing aspects of shape—than for the first two. Transfer trials in which pairs of cues were confronted with each other likewise showed contour vs. surface to be the most potent and round vs. sharp and symmetry to be the least potent cues at this general age level.

That a preferential response to the contour vs. solid-form cue as opposed to shape represents a relatively primitive stage of perceptual development is further indicated by an investigation by Knoblauch (1934), comparing mentally deficient *S*s, aged 7 to 21, with normal children, aged four and one-half to eight and one-half. Following an initial discrimination between a solid circle and various angular forms, the mentally deficient group was much more likely than the normals to choose solid figures, regardless of shape, in transfer trials, while the normals tended to respond on the basis of the round vs. angular shape of the stimuli.

Preferential choices on transfer trials of a stimulus-equivalence experiment are of course of equivocal significance as regards *S*'s *perception* of the various aspects of the stimuli. Yet the genetic primacy of the solid- vs. outline-form cue is of considerable interest, particularly in relation to Hebb's theory of perceptual development (1949), which postulates that segregation of figure from ground is present at birth, but that the perception of form requires, at least in

the human, a fairly extensive period of sensory learning.

The role of spatial orientation in the discrimination of form. As noted above, Ling (1941) found that the discrimination of his infants was little affected by changes in the orientation of the stimuli; this was true also in Gellermann's (1933) work with slightly older children. As a matter of fact, this seems part of a general tendency of young children to ignore spatial orientation in their response to form, which was observed many years ago by Stern (1909), both in the drawings of children and in their identification of familiar objects.

There are really two aspects to this problem: a positive and a negative one. The first refers to an actual superiority of children over adults in recognizing stimuli under changes in orientation, at least relative to their performance under normal conditions. This superiority has been found both for the recognition of inverted pictures (Mouchy, quoted by Köhler, 1940, p. 20f) and of nonsense words turned through 90° (Oetjen, 1916). The results of Newhall (1937), showing that the performance of three- to five-year-old children in a visual acuity task is not influenced by either right-left or up-down reversals, point in the same direction.

The other side of this coin is that young children experience great difficulty in discriminating between forms which are mirror-images or up-down reversals of one another. This is a well-known fact to students of the early development of reading skills (cf. Davidson, 1934, 1935; Vernon, 1957). Rice's (1930) study, sometimes cited in this connection, is somewhat equivocal, since it only showed a lack of spontaneous verbalization of the inversion of a figure in a matching task on the part of the younger children (four to five years). Similarly ambiguous is Burkhardt's (1925) finding of a sharp decrease between six and eight years in the incidence of orientational changes in the child's reproductions of a set of simple drawings. But there is more compelling evidence in an unpublished thesis by Newson (quoted by Vernon, 1957, p. 25), indicating an outright inability of five-year-old children to detect the difference between a shape and its mirror image, even when it is pointed out. Upside-down reversals caused somewhat less difficulty for these children than mirror-image reversals—a point most recently confirmed in a discrimination-learning experiment by Rudel (1959), with three- to five-year-old children.

The work considered thus far has dealt in the main with simple geometric forms, or individual familiar objects. A paper by Hunton (1955) throws an interesting sidelight on this question, for this author reports that in the description of complex meaningful pictures the inversion of the stimulus has the effect, up to age seven, of interfering with the perception of the *relationship* between the objects.

It may be pointed out in conclusion that this evidence on the weak role of the spatial orientation in young children seems to support the views of Würsten and Piaget, advanced to explain the developmental changes in the horizontal-vertical and other related illusions (cf. above). The evidence for improvement in spatial localization, reviewed in an earlier section, also fits rather well into this general picture.

Gestalt principles. Followers of the Gestalt school of perception have, in general, assumed an innate, intrinsic determination of their principles of

organization in the perception of form (cf. Zuckerman & Rock, 1957), and have accordingly not concerned themselves very intensively with the study of developmental changes. Piaget, among others, has taken exception to this view, arguing that the full-fledged operation of these principles is achieved only as the end-result of a prolonged developmental process. One particular study in which he has attempted to substantiate this claim (Piaget, Maire & Privat, 1954) consists of a comparison of the ordinary Müller-Lyer illusion, in which the two figures are shown superimposed vertically, with that resulting when the ends of the arrows are joined by vertical lines, so as to form a square in conjunction with the main lines of the Müller-Lyer figures (cf. Fig. 3). As one might expect, the connecting lines greatly reduce the illusion, in accordance with the stabilizing properties of a "good figure" such as the square. This effect is, however, *relatively* smallest for the youngest children, and greatest in adults, thus demonstrating, according to the authors, the increasing efficacy with age of these stabilizing properties. But the fact that the *absolute* amount of the difference between the illusions under the two conditions decreases with age somewhat weakens the force of this argument. On the other hand, the decrease with age in both inter- and intrasubject variability found for the modified figure, which contrasts with the increase in variability with age found for the ordinary Müller-Lyer figure, does indicate an increase in consistency which would support Piaget's point.

For some of the other principles of organization—that of continuation in particular—developmental changes are similarly demonstrable. A major

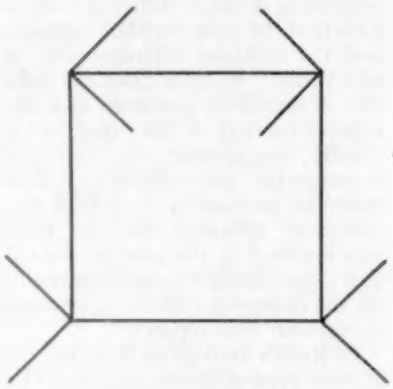


FIG. 3. Modification of Müller-Lyer figure, showing role of "good figure."

source of evidence on this problem is the investigation by Rush (1937), who has solved the difficult problem of quantification in this area with notable success (besides undoubtedly establishing a record for this type of experimentation as regards the total number of *Ss* tested: 8589!). Thus, in order to assess the strength of the *proximity* factor, this author presented her *Ss* with matrices of dots, varying the spacing of rows or columns and asking them to indicate by straight lines the direction in which they perceived the dots. No appreciable age changes emerged with respect to the role of this factor between the first grade and high-school levels—a finding which is in agreement with that noted by Schroff (1928), on the basis of rather less impressive data.

More interesting yet in Rush's study is her confrontation of various factors with each other. The author opposed the proximity factor to that of *continuation* (manipulated in terms of the number of dots in the rows and columns of the matrix), as well as to that of *similarity* (manipulated by

presenting elements differing in shape such that the rows were homogeneous and the columns heterogeneous, or vice versa). In both cases the influence of proximity decreased with age, relative to that of the other factor. Finally, juxtaposing the factors of continuation and similarity, Rush found an increase in the role of continuation, although this age trend was reversed in the case of the college group (similar anomalous results for the college-age *Ss* are in evidence throughout this report).

Of Rush's findings perhaps the one of most general significance is the increase in the effect of continuation with age, at least up to high school. Bearing in mind that this factor is represented here through discontinuous stimulus patterns—i.e., rows and columns of discrete dots and the like—it is tempting to relate this trend to the increase with age in the ability to achieve closure on the basis of incomplete figures, as shown in developmental studies on the recognition of pictured objects from partial cues (Brenner, 1957; Gollin, 1959a; Mooney, 1957; Schober & Schober, 1919; Van der Torren, 1907). (The normative data obtained by Street [1931] for his "Gestalt Completion" test from third-grade, sixth-grade, and high-school children fail to show appreciable age differences, possibly due to the relatively small number of stimuli employed.)

We might also mention in this connection Gollin's (1959b) work on the discrimination of tactual stimuli formed by patterns of tacks. This task is not only substantially more difficult for eight- to 10-year-old children than for adults, but the children are handicapped to a much greater extent by the presence of "noise," i.e., background stimuli not forming part of the patterns, in the field of touch.

If continuation in discontinuous or incomplete patterns becomes thus fully effective only at a rather late age, the role of this factor in the perception of uninterrupted line-drawings is on the contrary very strongly developed in young children (Ghent, 1956; Piaget & von Albertini, 1954). This difference between continuous and discontinuous stimuli is shown clearly in the study by Piaget and and von Albertini (1954); four- to five-year-old children experienced virtually no difficulty in tracing the outlines of forms when these overlapped with other forms. However, when these same forms were presented individually, but as dashed patterns, children up to the age of six could generally not recognize them; success in tracing these dashed forms superimposed as in the overlapping figures did not come until a still later age (between seven and nine years). Similarly, the completion of a partially mutilated geometrical figure also develops relatively late (i.e., about seven years) according to this study, these gaps proving especially difficult where oblique lines or angles are involved. Here immature completions, instead of conforming to principles of simplicity, tended frequently in the direction of "empirical," i.e., familiar, Gestalten.

These results thus suggest the importance of distinguishing between the role of closure in the perception of incomplete stimuli and the role of "good continuation" when continuous line patterns are involved. With respect to the latter, Ghent (1956) has confirmed its effectiveness in children as young as four years in the recognition of overlapping forms. On the other hand, in a second portion of Ghent's study, the recognition of embedded figures of the Gottschaldt type proved much more difficult for young children, giving rise to a

marked improvement between 4 and 13 years. This improvement is corroborated by other developmental studies of the perception of embedded figures (Gollin, 1956; Witkin et al., 1954), as well as by the similar results found by Heiss (1930) with respect to the ability to detect an angular shape when it forms part of a continuous outline.

The evidence reviewed here suggests, then, that young children have a very strong tendency to follow what may be loosely described as a "line of least resistance," i.e., a continuous linear path, in their perception of complex linear forms. This tendency clearly serves them in good stead in the recognition of overlapping forms, while it interferes with the isolation of component parts of a total outline (as in Heiss's study), and more particularly with the perception of angular forms competing with the continuous line pattern, as in the Gottschaldt figures. In this connection a study of developmental changes in eye movements in the visual exploration of such figures should prove revealing. At the same time the application of information-theory principles (cf. Attneave, 1954) suggests itself as the means for a more systematic attack in the study of these developmental changes.

Whole vs. part perception. Some of the problems just discussed, notably in the consideration of embedded and overlapping figures, are closely related to the question of the perception of part as against whole aspects of complex forms. Studies of age changes on the Rorschach represent our main source of information concerning the developmental processes involved here. They generally agree in finding a change from a preponderance of global percepts, determined by an unstructured perception of the whole in early childhood, through a phase in middle childhood in which large and small details predominate, to an eventual more integrated perception of the whole in the adult (Ames, Learned, Metraux, & Walker, 1953; Dworetzki, 1939; Hemmendinger, 1953).

In view of the lack of control over either the stimulus or the response in the Rorschach test, and the consequent equivocal significance for perceptual processes of the *S*'s interpretations of these blots, it is of some importance that this formulation of the developmental changes in part-whole perception has been substantially confirmed in several other studies which permitted a more systematic confrontation of part and whole aspects. Foremost among these is the investigation of Dworetzki (1939) which, besides the Rorschach cards, employed meaningful designs made up of individual parts which themselves represented meaningful objects. Here the trend in the *S*s' identifications of the designs paralleled rather closely that found in the Rorschach studies: the youngest children responded merely to the design-as-a-whole, while somewhat older children responded mainly to the component parts; adolescents and adults, finally, included both of these aspects in their descriptions of the stimuli.

A further study by Selinka (1939) represents perhaps a still more adequate approach to the question of the integration of the parts in the perception of the whole. This author constructed a task similar to the Kohs Block-Design test, in which a particular circular pattern shown to *S* had to be reproduced by means of four blocks, each of which contained a quarter of a circle made up of different kinds of patterns on the different sides. Thus the extent to which the pattern of the model was

taken into account in the arrangement of the blocks showed the child's attention to detail in his reproduction of the total circular outline. This task showed a considerable improvement from the kindergarten level, where the component patterns were usually ignored, through successive school-age groups up to 10 years. Acutally, the major portion of this improvement was from the seven- to the eight-year-old level, suggesting the possible influence of specific experience in school.

Various other studies (Segers, 1926; Smith, 1914) present further evidence regarding the difficulty experienced by young children in perceiving small details in complex but meaningful wholes. However, when the whole does not itself represent a strongly competing aspect of the stimulus, even young children can respond to details. This is shown in an experiment by Van der Torren (1907), who presented children with drawings of familiar objects at progressively increasing levels of completeness; at each trial they not only tried to identify the object, but also to point out the parts in the picture that had been added since the previous exposure. In line with the abovementioned age trends for incomplete figures, the identifications improved with age (i.e., in terms of the degree of completion required). Yet the recognition of the added details was almost perfect from the youngest age level (four years) on. Schober and Schober (1919), however, replicated this study with the single modification that all presentations of the drawings contained the complete outline of the object, so that the successive exposures differed only in the internal details. The result was that the recognition of the detail added to each exposure was greatly reduced for the younger children, increasing from 36% to 88% between the ages of four and eight years. Thus the presence of the completed outline of the object seemed in effect to interfere with the young child's perception of the parts of the figures.

Meili (1931) has taken these findings, together with related evidence, as indicating the young child's inability to attend to *both* the part and the whole at the same time. Which he will attend to will depend, according to this author, on the degree of structure present in the stimulus: for strong or simple wholes, the whole will be perceived, whereas for weak or complex wholes there will be a concentration on detail. Although this formulation represents a promising advance over the overgeneralized statement that young children are unable to perceive details in a whole, it is obviously of limited usefulness, pending an operational definition of strong vs. weak wholes and a determination of the interaction of this variable with that of complexity. Conceivably these two factors can be reduced to those of continuity of outline and degree of internal redundancy—a conceptualization which would bring this problem, together with that discussed in the previous subsection, into the analytic framework of information theory.

Number

The role of configurational factors in the identification of number, under exposure times too short to permit counting, is the subject of experiments by Freeman (1910) and Biemüller (1930). Freeman showed groups of from 3 to 12 dots arranged in a variety of configurations, as well as in vertical and horizontal rows. In general, the accuracy of the children (6 to 14 years old) started to lag be-

hind that of adults for groups of dots greater than five, with the exception of complete patterns (e.g., a 3×3 matrix to represent 9), which the children in particular found easier than corresponding incomplete patterns. There were, however, only 14 *S*s in the children's group, thus ruling out the determination of progressive age changes.

Biemüller (1930) limited himself to linear patterns of from 3 to 10 marbles exposed for ca. 1 sec. They were spaced either uniformly, or with gaps between some of the marbles, so as to break the row up in various ways, some rhythmical, some symmetrical about the middle of the row and some irregular. The added redundancy in the rhythmical or symmetrical patterns proved relatively ineffective in improving accuracy of the estimates below adolescence. There was, in fact, a marked tendency on the part of the younger children to ignore these groupings in their reproduction of the perceived patterns: rhythmical and especially symmetrical patterns were frequently turned into uniform, unpatterned rows. This study thus seems to indicate that the young child is to a degree handicapped in handling information with regard to patterning in discontinuous stimuli, which may be related to the relatively minor role of continuation at the younger age levels in Rush's (1937) study. As for the apparently conflicting results of Freeman, it should be noted that they refer to patterns of matrices, in which the configurational element is presumably much stronger; furthermore, the sampling is much less adequate and comprehensive than in Biemüller's experiment.

Lastly we may cite two studies which will serve as a convenient bridge to the following sections on time-bound phenomena. Judd (1927) presented series of from two to seven stimuli in temporal succession, in intervals of one second, either in the form of tones or of visual flashes. For both modalities there was a regular and very substantial improvement from the first to the sixth grades in the *S*'s accuracy of identification of the numbers. Interestingly enough, at all levels performance was poorer for the flashes than for the sounds; there was some indication, moreover, that this difference between modalities increased with age.

The second study, by Stambak (1951), although not dealing with number as such, is closely related to Biemüller's work on patterning, but over a temporal rather than a spatial dimension. In this experiment children between 8 and 15 years were asked to reproduce a series of 21 heard rhythmic patterns, made up of three to eight elements in one to five subgroups. This kind of temporal integration again improved with age, but there was almost perfect consistency in the order of difficulty of the patterns at the different age levels. The major determinant was the total number of elements in the pattern, but for the larger numbers the number of subgroups was an important contributing factor.

Movement

Apparent movement. Several studies on developmental changes in apparent movement have been reported (Brenner, 1957; Gantenbein, 1952; Meili & Tobler, 1931); all agree in finding a decrease in susceptibility to the perception of such movement with increasing age. The most thorough of the investigations on this problem is that by Gantenbein (1952), who has verified this age trend across a variety of conditions

of stimulus exposure, systematically varying distance, temporal interval, exposure time and intensity, with fairly consistent results, i.e., an increase with age in the threshold between succession and movement. The same applied to the threshold between movement and simultaneity, but here the age differences were much less pronounced. This was apparently due in large measure to the technical difficulty of producing an appearance of simultaneity in the adult (frequently no upper threshold could be determined). Indeed, in Brenner's (1957) study, just the reverse held true: age differences were greater for the movement-simultaneity threshold, simultaneity being perceived at lower speeds by the younger children.

Meili and Tobler (1931) attribute children's greater susceptibility to apparent movement to a lesser spatial stability of the stimulus in childhood. They discount an explanation in terms of a greater tendency for fusion in children, showing that the fusion threshold for a Talbot disc in a five-year-old group is essentially equivalent to that of an adult group. While the number of *S*s on which this comparison was based was rather small, several other experiments on the CFF confirm this finding (Hartmann, 1934; Miller, 1942).

Brenner (1957), on the basis of a somewhat mistaken interpretation of the argument proposed by Meili and Tobler, suggests that according to their explanation of the age changes in apparent movement this phenomenon should be correlated with speed of closure. This she finds not to be so: there is, on the contrary, an inverse correlation between scores on a Street-Gestalt test and the threshold for apparent movement. But there seems to be little reason to expect a relationship between closure, in the recognition of meaningful pictures from partial cues, and the stimulus displacement involved in apparent movement.

A phenomenon undoubtedly related to apparent movement is the subject of a study by Piaget and Lambercier (1951b): a square rotates along an eccentric circular path; as the speed of rotation increases, the rotating square changes first to a stationary cross and subsequently to two superimposed crosses, an inner and an outer one. Here again, as age increases progressively higher speeds are required for the perception of the single and double cross.

Movement in complex patterns. The perception of the *direction* of movement of a number of stimuli simultaneously displaced forms the topic of an interesting experiment by Rey (1954). A column of five dots is shown which can be displaced either to the right or left or remain stationary, each independently of the others. Various patterns of movement can thus be produced. The *S*'s task is to specify (from memory) which dots moved on any given trial, and in what direction. Where two or more stimuli moved in opposite directions, the correct identification of the direction of movement turned out to be particularly difficult for the youngest children (ages four to six): apparently they are unable to coordinate the left-right dimension with the vertical position of the dot. Furthermore, even if errors of direction are ignored, correct identification of the dots displaced on a given trial steadily decreased as the number of dots increased, but to a much larger extent in young as compared with older children.

Speed. In his unpublished thesis Edgren (1953) included tests of the

relative speed of two moving objects under a variety of conditions, as originally studied by J. F. Brown (1931) in adults. Age differences were, in general, either nonexistent or very irregular. The only suggestion of an age trend was observed in a rather difficult type of judgment in which the standard and variable were moving at right angles to each other.

Introducing the dimension of meaning into their stimuli, Wapner and Werner (1957) found much more striking developmental changes in perception of relative speed. They prepared a set of pairs of outline drawings of familiar objects, one in each pair representing the object in a static position and the other showing it in a position suggesting motion. With both objects exposed simultaneously on moving belts, the *S* had to adjust the speed of the latter (the variable) to that of the "static" object (the standard). For adults the PSE virtually coincided with the POE; not so for the children, who, to an extent progressively decreasing with age, overestimated the speed of the variable; thus for them the speed of movement was affected by the dynamic qualities suggested in the drawing. (In this experiment, as in many others, the possibility of a systematic bias due to the use of one type of stimulus as standard and the other as variable must not be overlooked; it is at least conceivable that *S*s might tend to underestimate a constant as compared to a variable speed, and that this effect might interact with age. As noted previously, this whole problem of the "error of the standard" is badly in need of fuller and more systematic treatment.)

Causality. Two developmental investigations of the perception of causality, based on the classical work of Michotte (1946), should be included here, since this phenomenon is clearly an aspect of motion perception. Olum (1956) investigated three of Michotte's conditions, varying the relative speeds of the two stimuli involved. Whereas most adults perceived the relationship between the two stimuli either in terms of a releasing effect (for one combination of speeds) or of a pushing effect (for the other two), children's responses were much more variable for any given situation. In particular, where the speed of the two stimuli was identical or closely similar, a number of children gave responses without any causal implications, such as passing, crossing, and even "tunnelling." It is of course difficult to decide whether the age changes observed here reflect perceptual differences or rather shifts in verbal habits in the description of a perceived relationship between two moving objects. One result appears, however, to have a well-founded perceptual basis: where the two speeds were in the ratio of 1:30, a large number of children perceived them as mutually approaching—a finding which Olum quite plausibly attributes to the intervention of stroboscopic movement between the stimuli (which, as we have seen, occurs more readily in young children).

Piaget and Lambercier (1958), in the most recent of the studies in the series from the Geneva laboratories, report on the same phenomenon; they likewise find noncausal types of responses in seven- to nine-year-old children, where the speeds of the two stimuli were in the ratio of 1 to 1. In fact, as late as 13 years, half of the *S*s gave such responses.

A variable more thoroughly examined in this study is the spatiotemporal interval between the stim-

uli at the point where the second starts to move. The younger children perceive a causal relationship between the stimuli only as long as they perceive a physical contact between them, while for the adults the causal effect can operate across a certain gap. On the other hand children perceive physical contact much more readily in the absence of actual contact, i.e., across much larger spatio-temporal intervals. It appears plausible to relate this finding again to their greater susceptibility to apparent movement, both phenomena demonstrating the lesser spatial stability of the stimulus for the younger child, as argued by Meili and Tobler (1931).

Time

Although the problem of time perception would seem to be intimately related to that of velocity, this relationship does not emerge too clearly in the developmental literature on this subject, which has limited itself to judgments of either unfilled time intervals or of those filled by some static stimulus.

Fraisse (1948) tested children at ages 6, 8 and 10, as well as an adult group, by means of a reproduction technique, with time intervals of .5, 1, 5 and 20 sec., both filled and unfilled. The main age difference appeared in the judgments of the two longest empty intervals, which were grossly underestimated by the youngest group; at the same time their judgments for the 20-sec. interval were characterized by extraordinarily large variability. Only slight age changes were found with the short time intervals, or with the longer intervals after the age of eight. On the basis of these results, Fraisse suggests a distinction between time *perception* and time *estimation*, the former applying only to very short time intervals (e.g., 1 sec.). Thus this author sees the younger children as deficient primarily in their estimation of time.

The judgmental nature of the inferiority of younger children is borne out to some extent by a recent experiment by Smythe and Goldstone (1957). They report a steady decrease, between seven years and adolescence, in the error made in identifying a 1-sec. interval by a method of limits; however, following repeated exposure to this interval identified as one second, the children from eight years upward improved very markedly in these absolute-type judgments, so much so that the former age differences were wiped out.

Lastly an older study by Gilliland and Humphreys (1943) is of interest, since it compared judgments of time intervals from 9 to 180 sec., obtained by three different methods: estimation, production, and reproduction, in two age groups (fifth graders and college adults). Although one might expect a superiority of the reproduction method, which does not require an absolute scale of time, the actual difference observed in favor of this method was barely significant. Apparently, for intervals of this length, the mediation of some internalized time scale is essential even in a comparative judgment task; this would be in line with Fraisse's point of the nature of the judgment for these longer intervals. As for the differences between the two age groups in this study, the adults were more accurate throughout, but there was no interaction, either between age and time interval, or between age and method.

Perceptual Learning

Various studies cited in this review have given more or less incidental

evidence of practice effects; interestingly enough the extent of such effects is in general approximately proportional to the level of initial performance, and thus increases with age. This trend is seen in several of the studies from the Geneva Laboratories (Lambercier, 1946a, 1946b; Piaget & Denis-Prinzhorn, 1953; Piaget, Maire, & Privat, 1954) in which judgments for a baseline condition were made at the beginning and at the end of the experimental session; the improvement usually resulting may of course be due either to practice per se or to the effect of the interpolated conditions.

The studies on cutaneous localization (Dunford, 1930; Renshaw, 1930; Renshaw & Wherry, 1931; Renshaw, Wherry, & Newlin, 1930) all included extensive practice sessions under a single experimental condition, and thus provide more systematic and unambiguous information on learning effects in this type of task. Dunford found improvement with practice again closely proportional to the initial performance at each age level; similarly in Renshaw's (1930) study the children, who were initially superior to the adults, likewise benefited from their practice to a greater extent.

We may note also a related exploratory study by Gibson and Gibson (1955) on the recognition of nonsense designs: over a series of trials, children between 8½ and 11 years of age showed considerable improvement, while a six- to eight-year-old group gave very little evidence of learning. This difference seemed to be related to the *Ss*' readiness to verbalize the main stimulus dimensions on which these complex stimuli varied.

Perhaps the most interesting as well as the most thorough of the studies on the effects of practice is that by Noelting (in press) on the Müller-Lyer illusion, investigating changes in this illusion over a series of 20 trials. Noelting found a progressive increase with age in such changes, which tended to counteract the illusion. The youngest groups (ages five to seven) actually showed either no improvement or an increase of the illusion with practice. A polynomial trend analysis of the error curves, furthermore, disclosed that there was curvilinearity at all age levels: the curves went through a minimum (or maximum), followed by a slight reversal.

The interpretation of these age differences in perceptual learning must await a more adequate explanation of the nature of such learning effects, particularly since, occurring without benefit of external reinforcement, these cases of improvement with practice are not readily accounted for in associative terms (cf. Wohlwill, 1958). One plausible hypothesis might be that in at least some instances these practice effects are attributable to adaptation processes—cf. Köhler and Fishback's (1950) interpretation of changes in the Müller-Lyer illusion, in terms of satiation. The available evidence on developmental changes in adaptation (Carini & Carini, 1958; Giannitrapani, 1958) indicates, however, rather a tendency for adaptation to *decrease* with age.

Perceptual learning and perceptual development. A question relating to perceptual learning in a more general sense concerns its role in the developmental changes in perception which we have considered in this review. In what sense and to what extent are these changes interpretable as learning effects? Unfortunately, few if any developmental studies have achieved the kind of control over *Ss*' prior experience that would be required for an answer to this question.

What would seem to be needed is an attempt to manipulate perceptual experience, both through prolonged specific practice (as in the studies by Welch [1939a, 1939b]) and through exposure to different types of stimulus environments. There are obvious limits to the feasibility of this type of approach with children, but developmental psychology can surely profit here, as elsewhere, from work of this type carried out on subhuman species (e.g., Gibson & Walk, 1956).

Even without such experimental control, effects of learning are clearly demonstrable in at least one developmental change: the improvement in the tachistoscopic recognition of words and objects (Edgren, 1953; Forgays, 1953; Hoffmann, 1927). Here at least two experential factors may be specified: increasing familiarity with the particular objects and words exposed, and, in the case of the studies on word recognition by Forgays and Hoffmann, the perfection of the visuomotor habits involved in reading. The role of the latter factor is shown in the sharp rise in performance observed in both of these studies between the second and fourth grades of school, coinciding thus with intensified instruction in reading. It is further attested to by the gradual development of selective recognition for words exposed to the right of the fixation point, demonstrated in Forgay's experiment. It need hardly be pointed out, however, that the generalizability of specific learning effects of this type to perceptual development in other areas (e.g., illusions, constancies) remains very much open to question.

Discussion

In the developmental changes brought out in this review one may discern, it seems, three distinctive trends which appear to cut across a variety of content areas. This concluding section will be devoted to a brief discussion of these trends, in the hope that some of the theoretical and experimental issues raised by this review will emerge in clearer focus.

Assimilation and contrast. The work on the illusions indicates that assimilation tendencies decrease with age, while contrast effects, at least of the temporal variety, increase. Further age changes which are probably relevant in this connection are the decreases with age in errors of anticipation or starting-position effects, when the methods of limits or adjustment are used (e.g., Binet, 1895; Liebert & Rudel, 1959; Piaget & Lambercier, 1953; Wapner & Werner, 1957), as well as the increase in compensatory tendencies such as those encountered in judgments of verticality under conditions of body tilt. We will return to the significance of this trend for the general problem of perceptual development at the end of this section.

The spatial framework. The work on the role of spatial orientation in form discrimination gives clear evidence that the younger child generally fails to relate a stimulus to the spatial framework in which it appears. The sharp decrease with age in errors of spatial localization and in judgments of verticality (in the absence of conflicting kinesthetic cues), horizontality, and the like point in the same direction, as do the changes with respect to certain illusions, notably increases up to a certain point for certain forms of the horizontal-vertical illusion and possibly for the Schumann square as well. The similar trends observed in the case of projective-size judgments probably represent another aspect of the same

process. In most of the work on this question, however, the framework has been relatively remote (e.g., the walls of the room, the edges of the table) or altogether implicit; conversely, in the rod-and-frame test of Witkin et al. (1954) involving a frame in close spatial proximity to the stimulus being judged, the influence of the framework is strongest at the younger age levels. The actual role of the spatial framework in perceptual development needs thus to be clarified; one suggestion would be to enclose stimuli within a near-by framework which is itself in conflict with the more remote but also more constant one defined, e.g., by the room.

The role of redundancy in the stimulus. The age changes in the recognition of incomplete forms point to a third important dimension of developmental change. Compared to the adult the young child requires more redundancy in a pattern to perceive it correctly; thus both incomplete and very complex patterns will be difficult for him. By the same token, wherever a stimulus permits of several alternative modes of perception, that aspect involving the least amount of information will be the one responded to by the young child. This interpretation is consonant with the marked tendency of young children to follow along continuous linear paths in their perception of shape (e.g., in the embedded-figure task) and to focus on the general outline of a figure rather than on internal detail. It is consistent, furthermore, with the low potency in early childhood of form as a cue in matching and stimulus equivalence experiments, when it is in competition with spatially more redundant cues such as outline vs. solidity (Knoblauch, 1934; Rüssel, 1931) and color (Colby & Robertson, 1942). Lastly, consider a phenomenon of an entirely different order: the decreasing dependency with age on multiple cues in the constancies. Here again the younger child requires a greater amount of surplus information, in the sense of the simultaneous presence of a variety of cues which are usually correlated.

The problem of veridicality. The work on assimilation and contrast effects point to a problem of considerable importance for theories of perceptual development: do changes in perception as a function of age or learning necessarily tend towards veridicality, as Gibson and Gibson (1955) have argued? Although starting from rather different premises from the Gibsons', Piaget in his conception of perceptual development likewise emphasizes the increasing accuracy of perception, resulting from the intervention of regulatory mechanisms such as "decentrations" and "transports" (cf. Piaget, 1950). Piaget's studies on the Delboeuf and parallel-line illusions bear out this view, demonstrating decreases with age both in assimilation and in contrast. In Piaget's view, the main distinction which needs to be made is between spatial and temporal phenomena, since in the case of the latter the regulatory mechanisms are not given an opportunity to intervene; thus the increase in successive contrast with age in the Usnadze illusion (Piaget & Lambercier, 1944).

Operating within the framework of sensory-tonic theory (Werner & Wapner, 1952), on the other hand, Wapner and Werner (1957) have arrived at a formulation which largely eschews any reference to veridicality in the developmental trends which it hypothesizes; these are rather specified in terms of the changing

interplay of a system of forces which may operate as readily to increase error (as in the development of the compensatory mechanisms in judgments of verticality) as to reduce it (as in the decreasing effects of static-object stimulation, exemplified in the starting-position effects in the same situation). With respect to the illusions, Wapner and Werner similarly imply that assimilation effects decrease with age, with a concomitant increase in the strength of contrast effects, as shown in their results for the Titchener-circles illusion.

In the view of this writer, this problem represents one of the central unresolved questions in the area of perceptual development. Further and more thorough investigation of this question would provide us with the empirical basis for a more adequate conceptualization of the character and direction of developmental changes, and would thus shed light on the relative merits of the models of perception proposed by Piaget and by Wapner and Werner. At the same time the clarification of age changes in assimilation and contrast should contribute to a fuller understanding of these perceptual processes and of their dimensionality—i.e., whether assimilation and contrast do in fact represent two sides of a single psychological continuum, as suggested in the work on anchoring effects (Sherif, Taub & Hovland, 1958) and on leveling and sharpening (Holzman & Klein, 1954).

If we have tended here to emphasize the differences between the conceptions of perceptual development advanced by Piaget and his collaborators on the one hand and by Wapner and Werner on the other, we should note, in conclusion, an interesting point of similarity in their respective modes of approach to the study of developmental changes in perception. Both of these theories, however different in the types of constructs with which they deal, concern themselves in the first place with certain basic processes of perceptual judgment, as they operate in the adult. Developmental changes are incorporated into these systems in an integral manner, being conceptualized in terms which bear a direct relationship to the basic constructs of the respective theories. Such an approach appears well worth emulating in this field; it would in fact go far towards breaking down the barriers between experimental and differential psychology to which Cronbach (1957) has pointed. In this connection we might reflect on the view expressed by Lewin, that "problems of individual differences, of age levels, and of general laws are closely interwoven . . . general laws and individual differences are merely two aspects of one problem; they are mutually dependent on each other and the study of one cannot proceed without the study of the other" (Lewin, 1954, p. 921).

REFERENCES

AKISHIGE, Y. Experimentelle Untersuchungen über die Struktur des Wahrnehmungsraumes. Part II. *Mitt. jur.-liter. Fak. Kyushu-Univer.*, 1937, No. 4, 23–118.

AMES, LOUISE B., LEARNED, JANET, METRAUX, RUTH, & WALKER, R. Development of perception in the young child as observed in responses to the Rorschach test blots. *J. genet. Psychol.*, 1953, 82, 183–204.

ATTNEAVE, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, 61, 183–193.

BALDWIN, A. L. *Behavior and development in childhood.* New York: Dryden, 1955.

BEYRL, F. Über die Grössenauffassung bei Kindern. *Z. Psychol.*, 1926, 100, 344–371.

BIEMÜLLER, W. Wiedergabe der Gliederanzahl und Gliederungsform optischer Komplexe. *Neue psychol. Stud.*, 1930, 4, 161–283.

BINET, A. La mesure des illusions visuelles chez les enfants. *Rev. Philos.*, 1895, **40**, 11–25.

BRENNER, MAY W. The developmental study of apparent movement. *Quart. J. exp. Psychol.*, 1957, **9**, 169–174.

BROWN, J. F. The visual perception of velocity. *Psychol. Forsch.*, 1931, **14**, 199–232.

BRUNSWIK, E. Zur Entwicklung der Albedowahrnehmung. *Z. Psychol.*, 1929, **109**, 40–115.

BRUNSWIK, E. Distal focussing of perception: Size-constancy in a representative sample of situations. *Psychol. Monogr.*, 1944, **56** (1, Whole No. 254).

BRUNSWIK, E. *Perception and the representative design of psychological experiments.* Berkeley: Univer. California Press, 1956.

BURKHARDT, H. Veränderungen der Raumlage in Kinderzeichnungen. *Z. pädag. Psychol.*, 1925, **26**, 352–371.

BURZLAFF, W. Methodologische Beiträge zum Problem der Farbenkonstanz. *Z. Psychol.*, 1931, **119**, 177–235.

CARINI, L. P., & CARINI, PATRICIA F. Adaptation to a tilted rod in schizophrenic boys. *Amer. Psychologist*, 1958, **13**, 333–334. (Abstract)

CARR, H. A. *An introduction to space perception.* New York: Longmans, Green, 1935.

COHEN, W., HERSHKOWITZ, A., & CHODACK, MARJORIE. Size judgment at different distances as a function of age level. *Child Develpm.*, 1958, **29**, 473–479.

COLBY, M. G., & ROBERTSON, J. B. Genetic studies in abstraction. *J. comp. Psychol.*, 1942, **33**, 385–401.

COMALLI, P., WAPNER, S., & WERNER, H. Perception of verticality in middle and old age. *J. Psychol.*, 1959, **47**, 259–266.

CRONBACH, L. J. The two disciplines of scientific psychology. *Amer. Psychologist*, 1957, **12**, 671–684.

CRUIKSHANK, RUTH M. The development of visual size constancy in early infancy. *J. genet. Psychol.*, 1941, **58**, 327–351.

DAVIDSON, HELEN P. A study of reversals in young children. *J. genet. Psychol.*, 1934, **44**, 452–465.

DAVIDSON, HELEN P. A study of the confusing letters B, D, P and Q. *J. genet. Psychol.*, 1935, **47**, 458–468.

DENIS-PRINZHORN, MARIANNE. *La perception de la distance en fonction de l'âge.* Doctoral dissertation in progress, Univer. of Geneva, 1959.

DUKES, W. F. Ecological representativeness in studying perceptual size-constancy in childhood. *Amer. J. Psychol.*, 1951, **64**, 87–93.

DUNFORD, R. E. The genetic development of cutaneous localization. *J. genet. Psychol.*, 1930, **37**, 499–513.

DWORETZKI, GERTRUDE. Le test de Rorschach et l'évolution de la perception. *Arch. Psychol., Genève*, 1939, **27**, 233–396.

EDGREN, R. D. *A developmental study of motion perception, size constancy, recognition speed and judgment of verticality.* Unpublished doctoral dissertation, Stanford Univer., 1953.

ESTES, BETSY W. Judgment of size. Paper read at Soc. Res. Child Develpm., Bethesda, Md., March, 1959.

FORGAYS, D. G. The development of differential word recognition. *J. exp. Psychol.*, 1953, **45**, 165–168.

FRAISSE, P. Étude comparée de la perception et de l'estimation de la durée chez les enfants et les adultes. *Enfance*, 1948, **1**, 199–211.

FRAISSE, P., & VAUTREY, P. The influence of age, sex and specialized training on the vertical-horizontal illusion. *Quart. J. exp. Psychol.*, 1956, **8**, 114–120.

FRANK, H. Untersuchung über Sehgrössenkonstanz bei Kindern. *Psychol. Forsch.*, 1925, **7**, 137–145.

FRANK, H. Die Sehgrössenkonstanz bei Kindern. *Psychol. Forsch.*, 1927, **10**, 102–106.

FREEMAN, F. N. Untersuchungen über den Aufmerksamkeitsumfang und die Zahlauffassung bei Kindern und Erwachsenen. *Veröff. Inst. exp. Pädag. Psychol. Leipziger Lehrervereins*, 1910, **1**, 88–168.

GANTENBEIN, MARIA-MARTHA. Recherches sur le développement des perceptions: XIV. Recherches sur le développement de la perception du mouvement avec l'âge (Mouvement apparent dit stroboscopique). *Arch. Psychol., Genève*, 1952, **33**, 198–294.

GELLERMANN, L. W. Form discrimination in chimpanzees and two-year old children. *J. genet. Psychol.*, 1933, **42**, 1–50.

GESELL, A., ILG, FRANCES L., & BULLIS, G. E. *Vision: Its development in infant and child.* New York: Hoeber, 1949.

GHENT, LILA. Perception of overlapping and embedded figures by children of different ages. *Amer. J. Psychol.*, 1956, **69**, 575–587.

GIANNITRAPANI, D. Changes in adaptation to prolonged perceptual distortion: A developmental study. Unpublished doctoral dissertation, Clark Univer., 1958.

GIBSON, ELEANOR J., & WALK, R. D. The effect of prolonged exposure to visually presented patterns on learning to discriminate them. *J. comp. physiol. Psychol.*, 1956, **49**, 239–242.

GIBSON, J. J. *The perception of the visual world.* Boston: Houghton Mifflin, 1950.

GIBSON, J. J., & GIBSON, ELEANOR J. Perceptual learning: Differentiation or enrichment? *Psychol. Rev.*, 1955, **62**, 32–41.

GIERING, H. Das Augenmass bei Schulkindern. *Z. Psychol.*, 1905, **39**, 42–87.

GILLILAND, A. R., & HUMPHREYS, D. W. Age, sex, method and interval as variables in time estimation. *J. genet. Psychol.*, 1943, **63**, 123–130.

GOLDSTEIN, K., & SCHEERER, M. Abstract and concrete behavior: An experimental study with special tests. *Psychol. Monogr.*, 1941, **53**, (2, Whole No. 239).

GOLLIN, E. S. Some research problems for developmental psychology. *Child Develpm.*, 1956, **27**, 223–235.

GOLLIN, E. S. The developmental study of cognition. Paper read at Soc. Res. Child Developm., Bethesda, Md., March, 1959. (a)

GOLLIN, E. S. Tactual form recognition. Paper read at East. Psychol. Ass., Atlantic City, N. J., April, 1959. (b)

GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **63**, 141–154.

GRUBER, H. E. The relation of perceived size to perceived distance. *Amer. J. Psychol.*, 1954, **67**, 411–426.

HANFMANN, EUGENIA. Some experiments on spacial position as a factor in children's perception and reproduction of simple figures. *Psychol. Forsch.*, 1933, **17**, 319–329.

HARTMANN, G. W. Comparison of the flicker thresholds in children and adults. *Child Develpm.*, 1934, **5**, 122–126.

HARTMANN, G. W., & TRICHE, A. Differential susceptibility of children and adults to standard illusions. *J. genet. Psychol.*, 1933, **42**, 493–498.

HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.

HEISS, A. Zum Problem der isolierenden Abstraktion. *Neue psychol. Stud.*, 1930, **4**, 285–318.

HELSON, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297–313.

HEMMENDINGER, L. Perceptual organization and development as reflected in the structure of Rorschach test responses. *J. proj. Tech.*, 1953, **17**, 162–170.

HOFFMANN, J. Experimentell-psychologische Untersuchungen über Leseleistungen von Schulkindern. *Arch. ges. Psychol.*, 1927, **58**, 325–388.

HOLWAY, A. H., & BORING, E. G. Determinants of apparent visual size with distance invariant. *Amer. J. Psychol.*, 1941, **54**, 21–37.

HOLZMAN, P. S., & KLEIN, G. S. Cognition system-principles of leveling and sharpening: Individual differences in assimilation effects in visual time-error. *J. Psychol.*, 1954, **37**, 105–122.

HUNTON, VERA D. The recognition of inverted pictures by children. *J. genet. Psychol.*, 1955, **86**, 281–288.

ISHII, O. Studies on the depth perception on two-dimensional drawings: I. The role of structural factors Y and T. *Jap. J. Psychol.*, 1956, **27**, 83–84. (English abstract)

JOHNSON, BETH, & BECK, L. F. The development of space perception: I. Stereoscopic vision in preschool children. *J. genet. Psychol.*, 1941, **58**, 247–254.

JUDD, C. H. Psychological analysis of the fundamentals of arithmetic. *Suppl. educ. Monogr.*, 1927, No. 32.

KLIMPFINGER, SYLVIA. Die Entwicklung der Gestaltkonstanz vom Kind zum Erwachsenen. *Arch. ges. Psychol.*, 1933, **88**, 599–628.

KNOBLAUCH, ELISABETH. Vergleichende Untersuchungen zur optischen Auffassung hochgradig schwachsinniger und normaler Kinder. *Z. angew. Psychol.*, 1934, **47**, 305–375.

KÖHLER, W. *Dynamics in psychology.* New York: Liveright, 1940.

KÖHLER, W., & FISHBACK, JULIA. The destruction of the Müller-Lyer illusion in repeated trials: I. An examination of two theories. *J. exp. Psychol.*, 1950, **40**, 267–281.

LAMBERCIER, M. Recherches sur le développement des perceptions: VI. La constance des grandeurs en comparaisons sériales. *Arch. Psychol., Genève*, 1946, **31**, 1–204. (a)

LAMBERCIER, M. Recherches sur le développement des perceptions: VII. La configuration en profondeur dans la constance des grandeurs. *Arch. Psychol., Genève*, 1946, **31**, 287–323. (b)

LEIBOWITZ, H., & HEISEL, M. A. L'évolution de l'illusion de Ponzo en fonction de l'âge. *Arch. Psychol., Genève*, 1958, **36**, 328–331.

LEWIN, K. Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology.* (2nd ed.) New York: Wiley, 1954. Pp. 918–970.

LEYER, K. Tiefenwahrnehmung in den Entwicklungsphasen. *Z. Psychol.*, 1939, **146**, 229–279.

LIEBERT, R. S., & RUDEL, RITA G. Auditory localization and adaptation to body tilt: A developmental study. *Child Develpm.*, 1959, **30**, 81–90.

LING, BING-CHUNG. Form-discrimination as

a learning cue in infants. *Comp. psychol. Monogr.*, 1941, **17** (2, Whole No. 86).

MASSUCCO COSTA, ANGIOLA. Constanza percettiva di intervalli vuoti. *Arch. Psicol. Neurol. Psichiat.*, 1949, **10**, 377–388.

MEILI, R. Les perceptions des enfants et la psychologie de la Gestalt. *Arch. Psychol., Genève*, 1931, **23**, 25–44.

MEILI, R., & TOBLER, E. Les mouvements stroboscopiques chez l'enfant. *Arch. Psychol., Genève*, 1931, **23**, 131–156.

MICHOTTE, A. *La perception de la causalité*. Louvain: Vrin, 1946.

MILLER, V. L. The critical frequency limen for visual flicker in children between the ages of 6 and 18. *Genet. psychol. Monogr.*, 1942, **26**, 3–53.

MISUMI, J. Experimental studies on the development of visual size constancy in early infancy. *Bull. Fac. Lit. Kyushu Univer.*, 1951, No. **1**, 91–116.

MOONEY, C. M. Age in the development of closure ability in children. *Canad. J. Psychol.*, 1957, **11**, 219–226.

MUNN, N. L. *The evolution and growth of human behavior*. Boston: Houghton Mifflin, 1955.

NEWHALL, S. M. Identification by young children of differently oriented visual forms. *Child Develpm.*, 1937, **8**, 105–111.

NOELTING, G. L'apprentissage perceptif de l'illusion de Müller-Lyer chez l'enfant et l'adulte. *Arch. Psychol., Genève*, in press.

OETJEN, F. Die Bedeutung der Orientierung des Lesestoffes für das Lesen und der Orientierung von sinnlosen Formen für das Wiedererkennen derselben. *Z. Psychol.*, 1916, **72**, 321–355.

OHWAKI, SONOKO. On weight perception, especially the formation of Charpentier's illusion in children. *Jap. J. Psychol.*, 1953, **24**, 257. (English abstract)

OLUM, VIVIAN. Developmental differences in the perception of causality. *Amer. J. Psychol.*, 1956, **69**, 417–423.

PETERS, W. Die Entwicklung von Wahrnehmungsleistungen beim Kind. *Z. Psychol.*, 1927, **103**, 129–184.

PIAGET, J. Recherches sur le développement des perceptions: X. Les illusions relatives aux angles et à la longeur de leurs cotés. *Arch. Psychol., Genève*, 1949, **32**, 281–307.

PIAGET, J. *The psychology of intelligence*. New York: Harcourt, Brace, 1950.

PIAGET, J. Essai d'une nouvelle interprétation probabiliste des effets de centration, de la loi de Weber et de celle des centrations relatives. *Arch. Psychol., Genève*, 1955, **35**, 1–24.

PIAGET, J., & DENIS-PRINZHORN, MARIANNE. Recherches sur le développement des perceptions: XVI. L'estimation perceptive des cotés du rectangle. *Arch. Psychol., Genève*, 1953, **34**, 109–132.

PIAGET, J., & DENIS-PRINZHORN, MARIANNE. Recherches sur le développement des perceptions: XXI. L'illusion des quadrilatères partiellement superposés chez l'enfant et chez l'adulte. *Arch. Psychol., Genève*, 1954, **34**, 289–321.

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: II. La comparaison visuelle des hauteurs à distances variables dans le plan fronto-parallèle. *Arch. Psychol., Genève*, 1943, **29**, 173–253. (a)

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: III. Le problème de la comparaison visuelle en profondeur (constance de la grandeur) et l'erreur systématique de l'étalon. *Arch. Psychol., Genève*, 1943, **29**, 253–308. (b)

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: V. Essai sur un effet d'"Einstellung" souvenant au cours de perceptions visuelles successives (Effet Usnadze). *Arch. Psychol., Genève*, 1944, **30**, 139–196.

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: VIII. Transpositions perceptives et transitivité opératoire dans les comparaisons en profondeur. *Arch. Psychol., Genève*, 1946, **31**, 325–368.

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: XII. La comparaison des grandeurs projectives chez l'enfant et chez l'adulte. *Arch. Psychol., Genève*, 1951, **33**, 81–130. (a)

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: XIII. La perception d'un carré animé d'un mouvement de circumduction (Effet Auersperg et Buhrmeister). *Arch. Psychol., Genève*, 1951, **33**, 131–195. (b)

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: XV. La comparaison des différences de hauteur dans le plan fronto-parallèle. *Arch. Psychol., Genève*, 1953, **34**, 73–107.

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: XXIX. Grandeurs projectives et grandeurs réelles avec étalon éloigné. *Arch. Psychol., Genève*, 1956, **35**, 257–280. (a)

PIAGET, J., & LAMBERCIER, M. Recherchés sur le développement des perceptions: XXXI. Les comparaisons verticales à intervalles croissants. *Arch. Psychol., Genève*, 1956, **35**, 321–367. (b)

PIAGET, J., & LAMBERCIER, M. Recherches sur le développement des perceptions: XXXIII. La causalité perceptive visuelle chez l'enfant et chez l'adulte. *Arch. Psychol., Genève*, 1958, **36**, 77–201.

PIAGET, J., LAMBERCIER, M., BOESCH, E., & VON ALBERTINI, BARBARA. Recherches sur le développement des perceptions: I. Introduction à l'étude des perceptions chez l'enfant et analyse d'une illusion relative à la perception visuelle de cercles concentriques (Delboeuf). *Arch. Psychol., Genève*, 1942, **29**, 1–107.

PIAGET, J., MAIRE, F., & PRIVAT, F. Recherches sur le développement des perceptions: XVIII. La résistance des bonnes formes à l'illusion de Müller-Lyer. *Arch. Psychol., Genève*, 1954, **34**, 155–202.

PIAGET, J., & MORF, A. Recherches sur le développement des perceptions: XX. L'action des facteurs apatieux et temporels de centration dans l'estimation visuelle des longueurs. *Arch. Psychol., Genève*, 1954, **34**, 243–288.

PIAGET, J., & MORF, A. Recherches sur le développement des perceptions: XXX. Les comparaisons verticales à faible intervalle. *Arch. Psychol., Genève*, 1956, **35**, 289–319.

PIAGET, J., & MORF, A. Les isomorphismes partiels entre les structures logiques et les structures perceptives. In J. Piaget (Ed.), *Logique et perception*. Paris: Presses Univer. France, 1958. Pp. 49–116. (Études d'Épistémologie génétique, VI)

PIAGET, J., & OSTERRIETH, P. A. Recherches sur le développement des perceptions: XVII. L'évolution de l'illusion d'Oppel-Kundt en fonction de l'âge. *Arch. Psychol., Genève*, 1953, **34**, 1–38.

PIAGET, J., & PÈNE, FLORENCE. Recherches sur le développement des perceptions: XXV. Essai sur l'illusion de la médiane des angles en tant que mesure de l'illusion des angles. *Arch. Psychol., Genève*, 1955, **35**, 77–92.

PIAGET, J., & TAPONIER, SUZANNE. Recherches sur le développement des perceptions: XXXII. L'estimation des longueurs de deux droites horizontales et parallèles à extrémités décalées. *Arch. Psychol., Genève*, 1956, **35**, 369–400.

PIAGET, J., & VON ALBERTINI, BARBARA. Recherches sur le développement des perceptions: XI. L'illusion de Müller-Lyer. *Arch. Psychol., Genève*, 1950, **33**, 1–48.

PIAGET, J., & VON ALBERTINI, BARBARA S. Recherches sur le développement des perceptions: XIX. Observations sur la perception des bonnes formes chez l'enfant par actualisation des lignes virtuelles. *Arch. Psychol., Genève*, 1954, **34**, 203–242.

PIAGET, J., & VURPILLOT, ELIANE. Recherches sur le développement des perceptions: XXVII. La surestimation de la courbure des arcs de cercle. *Arch. Psychol., Genève*, 1956, **35**, 215–232.

PINTNER, R., & ANDERSON, MARGARET M. The Müller-Lyer illusion with children and adults. *J. exp. Psychol.*, 1916, **1**, 200–210.

RENSHAW, S. The errors of cutaneous localization and the effect of practice on the localizing movement in children and adults. *J. genet. Psychol.*, 1930, **28**, 223–238.

RENSHAW, S., & WHERRY, R. Studies on cutaneous localization: III. The age of onset of ocular dominance. *J. genet. Psychol.*, 1931, **29**, 493–496.

RENSHAW, S., WHERRY, R. J., & NEWLIN, J. C. Cutaneous localization in congenitally blind versus seeing children and adults. *J. genet. Psychol.*, 1930, **38**, 239–248.

REY, A. Contribution à l'étude de poids chez les anormaux. *Arch. Psychol., Genève*, 1930, **22**, 285–297.

REY, A. La perception d'un ensemble de déplacements. *Rev. int. Filmol.*, 1954, 75–92.

REY, A. La localisation dans le plan d'un stimulus visuel punctiforme. *Arch. Psychol., Genève*, 1955, **35**, 41–64.

RICE, CHARLOTTE. The orientation of plane figures as a factor in their perception by children. *Child Develpm.*, 1930, **1**, 111–143.

RIVERS, W. H. R. Observations on the senses of the Todas. *Brit. J. Psychol.*, 1905, **1**, 321–396.

RUDEL, RITA G. Discrimination of direction of line in children. Paper read at East. Psychol. Ass., Atlantic City, N. J., April, 1959.

RÜSSEL, A. Über Formauffassung zwei- bis fünfjähriger Kinder. *Neue psychol. Stud.*, 1931, **7**, 1–108.

RÜSSEL, A. Ein entwicklungspsychologischer Beitrag zur Theorie der geometrisch-optischen Täuschungen. *Arch. ges. Psychol.*, 1934, **91**, 289–304.

RUSH, GRACE P. Visual grouping in relation to age. *Arch. Psychol., N. Y.*, 1937, **31**, No. 217.

SAGARA, M., & OYAMA, T. Experimental studies on figural aftereffects in Japan. *Psychol. Bull.*, 1957, **54**, 327–338.

SANSTRÖM, C. I., & LUNDBERG, I. A genetic approach to sex differences in localization. *Acta psychol.*, 1956, **12**, 247–253.

SCHOBER, G., & SCHOBER, ADA. Über Bilderkennungs- und Unterscheidungsfähigkeit bei kleinen Kindern. *Beih. Z. angew. Psychol.*, 1919, **19**, 94–137.

SCHROFF, E. Über Gestaltauffassung bei Kindern im Alter von 6 bis 14 Jahren. *Psychol. Forsch.*, 1928, **11**, 235–266.

SEASHORE, C. E., & WILLIAMS, MABEL C. An illusion of length. *Psychol. Rev.*, 1900, **7**, 592–599.

SEGERS, J. E. Recherches sur la perception visuelle chez des enfants agés de 3 à 12 ans et leur application à l'éducation. *J. Psychol. norm. path.*, 1926, **23**, 608–636; 723–753.

SELINKA, R. Der Übergang von der ganzheitlichen zur analytischen Auffassung im Kindesalter. *Z. pädag. Psychol.*, 1939, **40**, 256–278.

SHERIF, M., TAUB, D., & HOVLAND, C. I. Assimilation and contrast effects of anchoring stimuli in judgments. *J. exp. Psychol.*, 1958, **55**, 150–155.

SLOCHOWER, MURIEL Z. Experiments on dimensional and figural problems in the clay and pencil reproductions of line figures by young children: I. Dimension. *J. genet. Psychol.*, 1946, **69**. 57–75.

SMITH, F. An experimental investigation of perception. *Brit. J. Psychol.*, 1914, **6**, 321–362.

SMYTHE, ELIZABETH J., & GOLDSTONE, S. The time sense: A normative, genetic study of the development of time perception. *Percept. mot. Skills*, 1957, **7**, 49–59.

SPITZ, H. H., & BLACKMAN, L. S. A comparison of mental retardates and normals on visual figural aftereffects and reversible figures. *J. abnorm. soc. Psychol.*, 1959, **58**, 105–110.

STAMBACK, MIRA. Le problème du rythme dans le développement de l'enfant et dans les dyslexies d'évolution. *Enfance*, 1951, **4**, 480–502.

STERN, W. Über verlagerte Raumformen. *Z. angew. Psychol.*, 1909, **2**, 498–526.

STEVENSON, H. W., & MCBEE, G. The learning of object and pattern discrimination by children. *J. comp. physiol. Psychol.*, 1958, **51**, 752–754.

STEET, R. F. *A Gestalt completion test.* New York: Bur. Publ., Teachers Coll., Columbia Univer., 1931.

TAMPIERI, G. Ricerca sperimentale sulle deformazioni sistematiche nel confronto fra grandezze visive. *Riv. Psicol.*, 1955, **49**, 4ff. (*Ann. psychol.*, 1957, **57**, 192)

THOULESS, R. H. Phenomenal regression to the real object. *Brit. J. Psychol.*, 1931, **21**, 339–359.

UPDEGRAFF, RUTH. The visual perception of distance in young children and adults: A comparative study. *U. Ia. Stud. Child Welf.*, 1930, **4**, No. 4.

VAN BIERVLIET, J. J. Nouvelles mesures des illusions visuelles chez les adultes et les enfants. *Rev. Philos.*, 1896, **41**, 169–181.

VAN DER TORREN, J. Über das Auffassungs- und Unterscheidungsvermögen für optische Bilder bei Kindern. *Z. angew. Psychol.*, 1907, **1**, 189–232.

VERNON, MARGARET D. *A further study of visual perception.* Cambridge: Cambridge Univer. Press, 1954.

VERNON, MARGARET D. *Backwardness in reading.* Cambridge: Cambridge Univer. Press, 1957.

VOLKELT, H. Fortschritte der experimentellen Kinderpsychologie. *Ber. IX. Kongr. exp. Psychol.*, 1926, 80–135.

VURPILLOT, ÉLIANE. L'influence de la signification du matériel sur l'illusion de Poggendorff. *Ann. psychol.*, 1957, **57**, 339–357.

WALK, R. D., & GIBSON, ELEANOR J. A study of visual perception in the human infant with a visual cliff. Paper read at East. Psychol. Ass., Atlantic City, N. J., April, 1959.

WALKER, W. Über die Adaptionsvorgänge der Jugendlichen und ihre Beziehung zu den Transformationserscheinungen. *Z. Psychol.*, 1927, **103**, 323–383.

WALTERS, SISTER ANNETTE. A genetic study of geometrical-optical illusions. *Genet. psychol. Monogr.*, 1942, **25**, 101–155.

WAPNER, S., & WERNER, H. *Perceptual development: An investigation within the framework of sensory-tonic field theory.* Worcester: Clark Univer. Press, 1957.

WEBER, C. O., & BICKNELL, N. The size-constancy phenomenon in stereoscopic space. *Amer. J. Psychol.*, 1935, **47**, 436–448.

WELCH, L. The development of discrimination of form and area. *J. Psychol.*, 1939, **7**, 37–54. (a)

WELCH, L. The development of size discrimination between the ages of 12 and 40 months. *J. genet. Psychol.*, 1939, **55**, 243–268. (b)

WERNER, H. *Comparative psychology of mental development.* (Rev. ed.) New York: Int. Univer. Press, 1957.

WERNER, H., & WAPNER, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324–338.

WINCH, W. H. The Vertical-Horizontal illusion in school children. *Brit. J. Psychol.*, 1907, **2**, 220–225.

WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality through perception: An experimental and clinical study.* New York: Harper, 1954.

WOHLWILL, J. F. The definition and analysis of perceptual learning. *Psychol. Rev.*, 1958, **65**, 283–295.

WÜRSTEN, H. Recherches sur le développement des perceptions: IX. L'évolution des comparaisons de longueurs de l'enfant à l'adulte avec variation d'angle entre la verticale et l'horizontale. *Arch. Psychol., Genève*, 1947, **32**, 1–144.

ZEIGLER, H. P., & LEIBOWITZ, H. Apparent visual size as a function of distance for children and adults. *Amer. J. Psychol.*, 1957, **70**, 106–109.

ZUCKERMAN, C. B., & ROCK, I. A reappraisal of the roles of past experience and innate organizing processes in visual perception. *Psychol. Bull.*, 1957, **54**, 269–296.

(Received June 18, 1959)

Psychological Bulletin
1960, Vol. 57, No. 4, 289-290

A NOTE ON PROJECTION

PHILIP H. CHASE
Veterans Administration Hospital, St. Cloud, Minnesota

Murstein and Pryer's article on the concept of projection (1959) surely reflects laudable motives, and their review of relevant research contains a number of astute observations. There appear to be, however, rather glaring faults in formulation, categorization and definition.

Let us agree that "if 'projection' means everything it means nothing" (Murray, 1951, p. 13). It is then difficult to see, except by Murstein and Pryer's final definition of projection (1959, p. 370), why they quoted Zilboorg's quote from the *Malleus Maleficarum* (Murstein & Pryer, 1959, p. 353). Hallucination, not projection, is the term usually applied when "devils stir up the inner perceptions" so "that they appear to be a new impression—from exterior things." Quotes from Murphy and Sears (Murstein & Pryer, 1959, p. 354) are interesting but will be seen to bear more relation to "New Look" perception than to projection. We shall also see that the admonition above had no apparent effect on their final definition of projection.

One can take issue with the authors' classification system on the following grounds. Unless we wish to emasculate the term "projection," it would seem appropriate to use it *only* when the process referred to involves the attribution of internal characteristics to some external person or objects. Inaccurate extension of the term may well have played a major role in the confusion recognized, but, nevertheless, added to by Murstein and Pryer.

Projection is defined as "the process of unwittingly attributing one's own traits, attitudes or subjective processes to others" (English & English, 1958). In relation to this definition the Murstein and Pryer classification of "attributive projection" is clearly redundant and relatively useless for purposes of classification. A more useful classification would result from considering the presumed purposes of the mechanism.

Two major categories are immediately obvious. We might term one type *defensive projection* and the other *predictive projection*. In the case of defensive projection the mechanism is seen to operate in defense of the ego. One's own unacceptable or denied characteristics are attributed to another. Murstein and Pryer's "classical projection" clearly falls into this category, and we may dispense with the rather loaded term "classical." In the case of predictive projection a problem in forecasting or describing the behavior or characteristics of others is the task of the individual. In the face of insufficient evidence about the objects of his predictions or forecasts, the *S* tends to attribute some of his own characteristics to the object. The comments cited by Murstein and Pryer to support their category of "attributive projection" all involve the predictive or forecasting aspect as the primary purpose of the mechanism.

One can see that Murstein and Pryer are on tenuous ground when in discussing their "attributive projection" they attempt to dispense with both the unconscious and the self-concept. They state that this category is not concerned with either concept. It can be easily argued that personality projection of any kind cannot occur without a self-concept

to project from, and that the self-concept and projective process may be characterized by a considerable range of consciousness although it is likely that defensive projection, at the time it occurs, is largely unconscious.

It is difficult to understand what prompted Murstein and Pryer to include a category of "autistic projection." Relying for definition again on the dictionary (English & English, 1958) we find that "autism" is defined as: "a tendency in one's thinking or perceiving to be regulated unduly by personal desires or needs, at the expense of regulation by objective reality—" (p. 59). Is it unreasonable to classify all personality projection as autistic? In defensive projection the need which clearly exerts such a strong effect is ego defense. In predictive projection the need is a function of the lack of data about the object. If we follow Murstein and Pryer we would have to define all need-determined perception as projection.

The logic that prompted Murstein and Pryer to categorize the rationalization of projection as a *type* of projection is difficult to understand. Rationalization is a mechanism in its own right. If the *S*, particularly if he has been using defensive projection, becomes vaguely aware of the irrationality of his projection and also in the service of the ego rationalizes his projection, this is not necessarily a type of projection.

It was a pleasure to have the research review made available to the reader, and it was hoped that as the end of the article was approached the authors would resolve some of the earlier semantic confusion criticized above. In their final definition, however, they were unable to escape the consequences of their earlier thinking. Their attempt to limit their definition to "perception or judgments having to do with the personality" (Murstein & Pryer, 1959, p. 370) was laudable even though earlier they had dealt so cavalierly with the self-concept. This reader was totally unprepared, however, for the statement that projection is "the manifestation of behavior by an individual which indicates some emotional value or need of the individual" (Murstein & Pryer, 1959, p. 370). To accept this definition would mean that, in spite of Murray's warning, we would have to include *as projection* at least the following short list: reaction formation, rationalization, repression, sublimation, suppression, compensation, displacement, identification, introjection, fixation, obsession, compulsion, phantasy, dissociation, withdrawal, conversion, substitution, love, hate, eating and defecating, although items such as the Babinski reflex may be exempt.

Let us still agree with Murstein and Pryer that "projection" is a mechanism needing considerable research effort. Let us hope with them that more precise and useful definition and classification can be achieved. Let us not, however, in our endeavors to achieve scientific status for a concept neglect the principles of logic and functional definition.

REFERENCES

ENGLISH, H. B., & ENGLISH, AVA C. *A comprehensive dictionary of psychological and psychoanalytical terms.* New York: Longmans, Green, 1958.

MURRAY, H. A. Foreword in H. H. Anderson & Gladys C. Anderson (Eds.), *An introduction to projective techniques.* New York: Prentice-Hall, 1951.

MURSTEIN, B. I., & PRYER, R. S. The concept of projection: A review. *Psychol. Bull.* 1959, 56, 5, 353–374.

(Received December 1, 1959)

Psychological Bulletin
1960, Vol. 57, No. 4, 291–314

IMPRINTING:
EMPIRICAL BASIS AND THEORETICAL SIGNIFICANCE[1]

HOWARD MOLTZ
Brooklyn College

Interest in the effects of early experience upon the behavior of adult animals has increased significantly in recent years as evidenced by investigations of problems ranging from the genesis of filial behavior in aves to the development of tactile discrimination in primates. In general, this increase in interest has been due to a growing awareness that analysis of the role of early experience is of fundamental importance in understanding many aspects of adult behavior. Witness the amount of research that has been concerned with the influence of neonatal sensory deprivation on functions such as form discriminations, interocular transfer, and problem solving capacity. Consider also the number of studies designed to investigate the relation between early social experience and the sexual and parental behavior of the adult organism.

Although these studies have called attention to the effects of variations in neonatal experience on the ontogeny of a wide variety of vertebrate response patterns, we still lack any clear understanding of the manner in which these effects are exerted (Beach & Jaynes, 1954). For example, which particular developmental processes are involved when the sensory environment of an animal is restricted such that it subsequently shows increased emotional excitement (Melzack & Thompson, 1956) as well as decreased responsiveness to pain? (Melzack & Scott, 1957). That early experience is influential in structuring adult behavior can scarcely be doubted, but the means by which this influence operates is, at best, only incompletely understood.

The results of a recent study by Fantz (1957) suggest that desire for increased understanding of the manner in which intraorganic processes and extrinsic stimulus conditions interact in ontogeny can, in some cases, be better served by employing species-specific responses as indicators of this interaction rather than behavior established through instrumental conditioning. A species-specific pattern suitable in this respect would be one that occurs with predictable regularity very early in the life of the organism and which, although stereotyped in its expression, is nevertheless liable to extensive modification by manipulation of the sensory environment. The phenomenon which Lorenz (1935) designated as "imprinting" appears to meet these requirements. As to the importance of studying imprinting for understanding the effects of early experience on the organization of behavior, Thorpe (1951) writes: "It needs, and would repay full and precise experimental investigation almost more than any other aspect of animal behaviour" (p. 256).

It is the purpose of the present paper to attempt a critical evaluation of the research literature concerned with imprinting and, on the

[1] Research presented in this paper was supported by Research Grants M-1732 and M-2417 from the National Institutes of Health, Public Health Service. The writer wishes to acknowledge with appreciation the many helpful suggestions offered by Evelyn Raskin.

basis of this evaluation, to suggest hypotheses for further empirical study.

Some Definitions of Imprinting

In 1873 Spalding reported that incubator hatched chicks tended to follow persistently the first moving object to which they were exposed. Heinroth (1910) subsequently called attention to this phenomenon when he reported that graylag geese can be made to respond to humans in filial fashion in preference to adults of their own species if they are exposed to humans just after hatching. Several years later, Lorenz (1935), concerned with analyzing the functional significance of stimuli involved in releasing and directing the social behavior of birds, extended the empirical basis of Heinroth's observations and provided a theoretical framework within which to interpret these observations. Lorenz emphasized that in certain species of birds a wide variety of either animate or inanimate objects can acquire, in the absence of any conventional reinforcing agent, the capacity to evoke certain aspects of behavior that are ordinarily directed toward members of the species.

Since the conditions under which an object can acquire this capacity, and the characteristics of the behavior thus evoked, were considered unique, Lorenz designated the process or mechanism involved by a special term—imprinting. To distinguish imprinting from associative learning, he emphasized the following aspects of the imprinting process:

1. Imprinting can occur only during a very definite period of the animal's life; a period of short duration during which the organism is assumed to be in a critical stage of physiological development. Thus if a graylag gosling, for example, is to respond to a human (or any other relatively large moving object) in filial fashion, initial exposure to the human must occur within several hours after hatching. The effect of this exposure is at first shown by the fact that the gosling persistently follows its "surrogate parent."

2. Once imprinting occurs its effect is irreversible. That is, the stimulus to which the animal is exposed during the brief critical period henceforth becomes either the preferred stimulus or the only stimulus toward which the following response will be directed. This preference, once it is established, remains extremely stable presumably throughout the life of the animal. Thus with reference to graylag geese, Lorenz (1957) states: "Once their instinctive social reactions are transposed to a human being, their behavior does not change in the least even if they are kept for years with other members of their own species and without human company" (p. 105). While this statement is perhaps too extreme to be taken as representative of Lorenz' views on the question of stability, there is no doubt that he believed imprinting produces modifications in behavior that were extraordinarily resistant to change.

3. Lorenz pointed out that many aspects of adult behavior that are not functional in the neonatal period during which imprinting is operative will nonetheless be directed toward the "imprinted object" when functional status is subsequently achieved. The fact that exposure to an object early in ontogeny can influence, in so specific a fashion, as yet undeveloped response patterns was considered characteristic of the imprinting process. Lorenz cited the case of adult shell parakeets that directed their

courtship activities toward humans in preference to available members of the species as a consequence of their having been "imprinted" to humans. This aberrant sexual behavior was held to be the result of imprinting and not of conditioning since, as Lorenz (1955 maintained, " . . . you cannot condition any not-yet-functioning response in the ordinary way" p. 209).

4. When a bird becomes "imprinted" to a particular stimulus it will readily transfer its following response as well as other social responses to all members of the class to which that stimulus belongs. In other words, imprinting was assumed to result in attachment not to the specific features of an object but to its general characteristics. Thus when a gosling, for example, is hatched under a muscovy duck, it subsequently follows not only its particular "foster parent" but all members of the species to which its "parent" belongs. Lorenz conceived of imprinting (as it operates in nature) as a method of acquiring a "consciousness of species"; a method that insures that social behavior will extend beyond the parent but not beyond the species.

Many contemporary ethologists would decline to draw as sharp a line between imprinting and associative learning as Lorenz had done earlier. Indeed, in a recent publication Lorenz (1955) maintained that imprinting and conditioning should perhaps be considered as continuous rather than as distinct processes. But while the attempt to distinguish in an absolute manner between imprinting and conditioning has been abandoned, there nevertheless remains an emphasis on the innate aspects of the imprinting process. Thorpe (1956), for example, conceives of imprinting as functioning to complete for the animal " . . . the processes initiated by its inherited constitution" (p. 118). He points out that many avian species at the time of hatching possess a neurosensory mechanism that is initially responsive not only to stimuli provided by the parent but to stimuli provided by a variety of moving objects that are quite unlike the parent. Since, as Thorpe assumes, the activation of this mechanism functions to release the following response, its lack of selectivity renders the social behavior of the animal potentially susceptible to control by almost any relatively large object that possesses the quality of motion. It is obvious that this condition, if it were to continue, could hardly prove biologically adaptive, so that early in ontogeny the animal must acquire a preference for a particular class of objects toward which to direct its social behavior. Exposure to any moving object during the critical period serves to establish this preference by making the releasing mechanism more selective so that henceforth it will be activated only by that object or by objects of similar configuration. Fabricius (1951a) and Ramsay (1951) also consider imprinting to be the process that increases the selectivity of the releasing mechanism, insuring that the social behavior of the animal will be evoked by a relatively limited range of stimuli.

Lorenz (1935) was the first to point out that there were marked differences among avian species with respect to their "imprintability" and was the first to ascribe these differences to genetically determined variations in the specificity of a hypothetical releasing mechanism. Lorenz (1957) emphasized that some birds can be imprinted only to members of the species since their

" . . . innate perceptory patterns are so highly differentiated" (p. 267) that the following response cannot be released by any other object. Thus curlews and great godwits, for example, cannot be made to direct their social behavior to such biologically inappropriate objects as a human or a cardboard box. There are many other avian species, however (e.g., graylag goose, and mallard duck), which can be imprinted readily to moving objects that differ greatly in size and shape. This is assumed to be due to the fact that in these species the structure of the releasing mechanism at the time of hatching is differentiated only to the extent that it can be activated by moving objects and not by stationary objects.

The conception of imprinting as a process that functions to increase the specificity of a hypothetical releasing mechanism carries with it the responsibility of specifying precisely the behavioral events to which imprinting is functionally related. The attempt to establish these relationships has involved the use of such disparate responses as following an object during the neonatal period and courting the same object in later life. Indeed, Thorpe has suggested that recognition of territory (Thorpe, 1945) and acquisition of a particular song through association with members of the species (Thorpe, 1956) should also be considered as the result of the imprinting process. But while such indices of imprinting have been the subject of much speculation, they have received little empirical examination.

Recent investigators, in their effort to study the imprinting phenomenon systematically, have tended to concentrate experimental attention almost entirely on the following response. The present paper will also deal primarily with the following response and, in addition, will use the term imprinting to denote a particular experimental operation and *not* a process or a mechanism. It is realized that by proceeding in this way the uniqueness and, indeed, the "flavor" of imprinting become somewhat "attenuated" but (as will be shown) such attenuation is believed necessary in view of the paucity of available empirical evidence. Thus, imprinting will be defined as the procedure of visually presenting to an animal a large moving object during the first several hours of its life under conditions that insure that the object is not associated with such conventional reinforcing agents as food and water. This procedure has been found to evoke a close following of the object in such precocial avian species as ducks, geese, coots, moorhens, and domestic fowl. However, imprinting as defined above, has *not* been unequivocally shown to induce following either in passerine birds or in mammals. While Lorenz reports that he was able to cause young jackdaws to respond to humans in filial fashion and while Hediger (1950), Collias (1956), and Gray (1958) speak of imprinting in buffaloes, sheep, and guinea pigs, their findings require experimental confirmation under more carefully controlled conditions before they can be generally accepted. To broaden the definition of imprinting so that the present discussion might encompass these species and perhaps, in addition, include a greater variety of behavior patterns is not likely to prove theoretically salutary. It need hardly be added that systematic significance is not necessarily precluded by declining to consider the entire spectrum of behavior that has been assumed (but not demonstrated) to be susceptible

to modification by the imprinting procedure.

Imprinting and the Critical Period

One reason that imprinting has generated so much interest is that the effectiveness of the procedure appears to be restricted to a specific period of development. It will be recalled that Lorenz maintained that if an object is subsequently to elicit following, *initial* exposure to that object must occur within a relatively brief period after hatching; a period during which the bird was assumed to be in a critical stage of physiological development. Recently, several investigators have attempted to obtain evidence regarding the range of the critical period. For example, Fabricius (1951b) reports that in tufted ducks following can be elicited only if initial exposure occurs before 12 hours of age, while in the mallard, readiness to follow a strange moving object was still apparent in some cases after exposure was delayed as long as 72 hours. Even in the mallard, however, the probability of the animal following a strange object was found to be inversely related to the age at which initial exposure occurred. Alley and Boyd (1950) have shown that for the coot imprinting is most likely to prove effective if it is administered within the first eight hours from the occurrence of hatching. Ramsay and Hess (1954) found that for both mallard ducklings and Bantam chicks, 13–16 hours after hatching appeared to constitute the critical period. They reported that as little as 10 minutes of exposure to a male model decoy during this time was sufficient to produce a strong approach response in a subsequent choice test that involved the familiar model and an unfamiliar female decoy. The incidence of correct choices was found to decrease sharply when initial exposure occurred either before 13 hours or after 16 hours.

While there is undoubtedly a critical or sensitive period during which exposure to a moving object is most effective with respect to inducing subsequent following, it is nonetheless possible to elicit the response in some cases even if initial exposure is delayed long past the period that has been designated as "critical" for a given species. Steven (1955), for example, reports that a lesser white fronted goose that was captured wild when between one and two weeks of age soon developed a strong following response to man. Under controlled laboratory conditions, Moltz and Rosenblum (unpublished) found that of 16 Peking ducks kept in isolation for the first three days of life, five showed a strong following response after exposure on the fourth day to either a moving cardboard box or a model decoy. However, when initial exposure occurred during the critical period (i.e., 8–10 hours after hatching) 14 of 16 birds displayed vigorous following. After establishing the fact that in domestic chicks the critical period appeared to be confined to within the first 24 hours after the occurrence of hatching, Jaynes (1957) was able to induce strong following of a cardboard cube in inexperienced chicks as old as six days by exposing them to the cube for several consecutive hours. While the original report by Spalding on the following of domestic chicks contains only a few procedural details, it is apparent that he observed close following in some birds on the third day of life after removing the opaque hoods which these animals had been wearing from the moment of hatching.

Hinde, Thorpe, and Vince (1956) reported that as initial exposure to a stimulus object was delayed, moorhens became progressively less likely to follow, while coots were not as markedly affected. Indeed, it was found possible to induce following in some coots even after a delay of six days. These authors suggested that the difference between the two species with respect to the age at which following can be established is due to the fact that the moorhen, unlike the coot, exhibits early in its development a tendency to flee from strange moving objects. The appearance of the flight response was considered responsible for limiting the critical period in moorhens to within several hours after hatching. Since the coot first manifests this response at a somewhat later stage of development, it shows greater flexibility with respect to the age at which a strange object can elicit following. Hinde (1955b) concluded that the critical period is not a property of an inferred imprinting process but a consequence of the appearance of response tendencies incompatible with following.

On the basis of these studies, it seems reasonable to speak of a critical period with respect to the organization of the following response as long as one does so in relative rather than absolute terms. For while there is undoubtedly a relatively brief stage in the ontogeny of precocial birds during which imprinting is most likely to induce following, it does not appear to be either as rigid or as absolute as Lorenz implied. The fact that in some cases following occurs after initial exposure has been delayed for several days, raises doubt concerning the validity of conceiving the critical period as being specifically a function of the hereditary constitution of the organism. To be sure, even if the critical period were less flexible than it appears to be, this lack of flexibility would not, in itself, provide support for the argument that the critical period is the product of a genetically determined condition that arises independently of early experience. An animal might be responsive to certain extrinsic stimuli at one particular stage of ontogeny and be totally unresponsive at both an earlier and later stage, not because of the unfolding of some "innate growth plan," but because of changes brought about through the progressive interaction between the developing organism and its sensory environment (Schneirla, 1956). It is in terms of this interaction that the attempt will be made subsequently to account for the critical period and for the initial appearance of the following response.

The "Irreversibility" of the Following Response

It will be recalled that Lorenz contended that imprinting produces irreversible modifications in behavior and that this is one of its distinguishing features. The display of abnormal sexual fixations by adult birds which earlier in life had been imprinted to man, or to other "heterospecific parents," has frequently been adduced in support of this contention. The circumstances under which this evidence has been obtained, however, leave much to be desired with respect to the exercise of sufficient experimental control. Hinde (1955b), for example, states that "No really adequate experiments seem to have been done here, however, for in all cases known to the writer the birds were continually in the presence of man throughout their pre-adult life so that the attachment of the sexual response to man, though perhaps in-

fluenced indirectly by the early experience, could have been due to a later learning process" (p. 21). With reference to these sexual fixations, Thorpe (1956) has also stated that " . . . since special precautions against reinforcement by subsequent conditioning were not taken, this must have played a large, perhaps very large, part in the later course and strength of the phenomenon" (p. 358). The only studies relevant to the problem of "irreversibility" that have been performed with reasonable precision are those that have dealt with the following response. It is these studies to which attention will be directed.

The statement that imprinting produces behavioral modifications that are irreversible can be interpreted in two ways: (*a*) once exposure to a particular object induces following, a markedly dissimilar object cannot do so, and (*b*) following a familiar object will continue to occur without any significant decrement in response strength. The first interpretation is concerned with the problem of stimulus generalization, while the second is concerned with the problem of response stability during successive presentations of the same object. For convenience, they will be treated independently.

Generalization

Only a few studies have been reported that are relevant to the problem of generalization and of these, none has attempted to vary systematically the generalization stimuli. The major concern has been whether animals that have followed a particular object will subsequently follow an unfamiliar object when the latter is presented alone or in combination with the familiar object. Since the relation between familiar and unfamiliar was not ordered along any meaningful continuum in these studies, no precise generalization function can be specified.

Fabricius and Boyd (1952–1953) exposed one group of mallard ducklings to a balloon and a second group to a brown box within the first 24 hours from the occurrence of hatching. After several days, during which time repeated presentation of the familiar object occurred, the birds were exposed to the unfamiliar object either alone or in combination with the familiar one. Fabricius and Boyd reported that several birds followed the unfamiliar object, but in all cases the "response was less intense." Unfortunately, no data were presented which would permit one to determine the relation between initial response strength and the magnitude of the generalization decrement.

Using coots and moorhens, Hinde, Thorpe, and Vince (1956) conducted a number of generalization tests, employing objects as different from the training object as a box moving along a wire differs from a man walking. The experimental procedure involved presentation of the generalization object either immediately following initial exposure or after a number of opportunities to respond to the training object had been given. The results showed that many birds followed the generalization model on the first occasion on which it was presented, although following was frequently hesitant and unstable. There is no doubt, however, that generalization occurred and that response to an unfamiliar model was a positive function of the strength with which the training model was followed. Hinde, Thorpe, and Vince concluded that imprinting does not "irreversibly modify" behavior, since following one object does not pre-

clude subsequent following of markedly dissimilar objects.

Jaynes (1956) has presented data on domestic chicks which agree in most respects with the data obtained from coots and moorhens. The procedure Jaynes employed consisted of exposing newly hatched chicks to one of two dissimilar cardboard objects (i.e., a green cube or a red cylinder) that moved irregularly about an alley. After following had been well established, the unfamiliar object was presented alone. In all cases a decrement in responding to the unfamiliar object was evident, although examination of the data left no doubt that considerable generalization occurred. On this basis Jaynes concluded that it is "clearly untrue" that imprinting produces irreversible effects if irreversibility is construed to mean a high degree of specificity.

Moltz and Stettner (unpublished) attempted to test for generalization using an apparatus and an experimental procedure similar to that employed by Jaynes. They exposed one group of Peking ducks for 25 minutes per day for three days to a model decoy that traveled about a 10-ft. alley. A second group was exposed to a green cardboard cube that described the same pattern of motion as the decoy. On either the fourth or fifth day one half of each group was presented with the object to which it had not been previously exposed. Behavior in relation to the unfamiliar object ranged from initial avoidance to close and persistent following. Of those birds that followed during the generalization test, none reached the level of responding that was in evidence prior to the test. It is rather surprising that the magnitude of the generalization decrement did not appear related in any consistent manner to initial response strength (rho $= -.07$). Jaynes also reported inability to predict magnitude of this decrement from knowledge of the strength with which the familiar object was previously followed. It will be recalled that, in this respect, the results of Hinde, Thorpe, and Vince are contradictory. The fact that the generalization decrement exhibited by coots and moorhens covaried with initial response level, while that exhibited by Peking ducks and domestic chicks did not, cannot be explained on the basis of available empirical evidence.

It is obvious that a systematic investigation of imprinting in relation to the generalization phenomenon is required. Future research might be directed toward trying to order stimuli along some meaningful continuum (such as size or shape) so that a more precise generalization function can be specified. In addition, the relation between variations in the imprinting procedure (e.g., duration of exposure, energy expended in following, etc.) and the subsequent ability of the bird to respond selectively to a familiar object also requires study. But perhaps before investigating any of the above mentioned parameters, the various procedures used to test for generalization should be subjected to experimental analysis, since presenting the familiar and unfamiliar objects simultaneously rather than successively has been found to exert different effects on the temporal course of the generalization decrement.

Response Stability

As was noted previously, there are two ways to interpret the statement that imprinting produces irreversible modifications in behavior. The first interpretation is relevant to the problem of stimulus generalization and

the contention that once exposure to a particular object induces following, subsequent following of dissimilar objects is precluded. The second interpretation is that following a familiar object will continue to occur without any significant decrement in response strength.

Before proceeding to discuss the problem of response stability, it should be recalled that one of the conditions specified in defining imprinting is that the object to which the bird is exposed initially is not to be associated with any conventional reinforcing agent. This condition should also be made to hold for all subsequent exposures. If it can be shown that following remains stable in the absence of any obvious reward, then the attempt to identify the factors involved in maintaining the response is likely to prove systematically significant. On the other hand, if following results in the receipt of food or water (or if the object followed acquires secondary reward value through association with such reinforcing events) the problem of response stability becomes trivial. Consequently, the analysis below will be restricted to an examination of those studies concerned with determining the functional course of the following response under conditions that appear to involve neither primary nor secondary reinforcement. However, it should be mentioned that in some cases it was difficult to determine the extent to which reinforcement had been involved. Many "imprinting studies" have been carried out under seminatural conditions and reports of these studies often do not supply sufficient detail concerning the method employed. Especially when the experimenter was the object followed (as in Lorenz' early work), it becomes pertinent to inquire whether the animals were fed by the experimenter and, if so, the extent to which he became "contaminated" with reinforcement. The studies that follow seem to be sufficiently free from such contamination.

Fabricius (1951a) reported that although he was able to establish a strong following response in tufted ducks, shovellers, and eiders, the following response "gradually diminished" beginning at about three weeks of age. Nice (1953) employed essentially the same procedure as Fabricius, and also found that in shovellers following began to wane at approximately the same age. Although in many respects Fabricius' study represented a pioneer effort, it was performed under conditions that left much to be desired with regard to the degree to which the bird's experience in the test situation was controlled. Under somewhat more adequately controlled test conditions, Fabricius and Boyd (1952–1953) found that mallard ducklings, within the first 10 days of age, ceased following a moving box to which they had previously shown strong "attachment." Fabricius and Boyd contended that this rapid decrement in response strength was in part a function of the fact that the animals were housed collectively. Why this should interfere with following was not made clear.

Following appears to be somewhat more stable in coots and moorhens, for Hinde, Thorpe, and Vince (1956) observed that a decrease in response strength first became evident after four to five weeks of testing in moorhens and not until seven to eight weeks in coots. There was no doubt, however, that even the intense following of the coot began to wane and that "by the end of the juvenile

phase was lost altogether" (Thorpe, 1956, p. 363).

Moltz and Rosenblum (1958a) attempted to determine the functional course of the following response in Peking ducks under laboratory conditions that remained constant over successive test trials. Each bird was tested for 25 minutes per day for a period of 15 days. The object was a cardboard cube that described a fixed pattern of motion about an enclosed alley. The results indicated that following tended to reach a maximum by the third trial but shortly thereafter a progressive decrement in the strength of the response occurred. Activity irrelevant to following began to appear during either the sixth or seventh trial and increased in frequency during each successive exposure. Toward the end of the test series, the majority of birds not only failed to follow but evinced no interest in the object as it passed about the alley. This behavior is to be contrasted with the strong and rather intense following that was evident earlier in the test series.

On the basis of the few available studies presented, there can be little doubt that a significant decrement in the strength of the following response occurs with repeated exposure to a familiar test object. Indeed, in most of the species observed the response proved quite unstable, "disappearing" within a relatively short time after it was established. Thus, there is no empirical justification for maintaining that imprinting "irreversibly modifies" behavior, if the index of this modification is considered to be the strength at which following is maintained. It is possible, of course, that imprinting does exert a lasting influence on avian behavior, but that the following response simply does not reflect this influence. However, since imprinting, *uncontaminated by reward learning*, has not been shown to affect directly any other response system of the adult animal, this possibility remains unexplored.

Anxiety and the Stability of the Following Response

The fact that a number of studies show that intense following exhibited during the early test trials begins to decrease, is puzzling in view of the conditions under which the response was evinced. The decrement is obviously not due to the withdrawal of reinforcement since, at least in the obvious sense of the term, no reinforcement was ever present. Nor does the decrement appear to be a function of any other change in experimental conditions. In the study reported by Moltz and Rosenblum, (1958a) for example, the animals were maintained in individual cages from the time of hatching and throughout the experiment the only *moving* stimuli to which any animal was exposed were provided by its own body and by the test object. Furthermore, the conditions under which the study was performed made it possible to insure that the stimulus situation did not change over successive trials. Despite these precautions, however, the strength of the following response decreased progressively during the last half of the test series. The mechanism mediating this decrement requires some discussion.

Several investigators have commented that, beginning at approximately 25 or 30 hours from the occurrence of hatching, precocial birds will frequently exhibit "anxiety" or "fear" in response to unfamiliar aspects of the environment. Ramsay and Hess (1954), for example, noted that "fear responses" were a char-

acteristic feature of mallard behavior and Fabricius (1951a) reported the same phenomenon in tufted ducks and eiders. Hinde, Thorpe, and Vince (1956) speak of the fleeing response in moorhens and Jaynes (1958a) found the same behavior in the domestic chick. The writer has also observed that the Peking duck responds in a similar manner to a wide variety of unfamiliar stimuli. It was noted, however, that this behavior was not present while the bird was following. Indeed, when in close proximity to the familiar test object, even the sudden introduction of a loud sound would sometimes fail to elicit any sign of fear. The impression was gained that following served to decrease anxiety and that the response would continue to occur only as long as anxiety continued to be aroused by the test situation.

Anxiety will be defined as an internal emotional state of the organism, this inferred state being indexed by such observable events as distress calls, startle behavior, defecation, etc. On the assumption that strength of following is some positive function of this test-aroused anxiety, it appears reasonable to expect that any experimental procedure designed to reduce anxiety would reduce the strength of the following response. Previous research (Moltz, 1954) has indicated that forcing an animal to remain in a situation in which anxiety has developed will either decrease or eliminate that anxiety. Accordingly, we can now formulate the following specific hypothesis: exposing ducklings to a test situation in the absence of the object to which a strong following response had been previously attached will subsequently produce a decrement in the strength of that response.

Evidence supporting this hypothesis was recently reported by Moltz and Rosenblum (1956b). They exposed Peking ducks for 25 minutes per day for three days to a cardboard cube that moved about an enclosed alley. Those birds which showed evidence of strong following were retained for further study and assigned to either an experimental or a control treatment. Beginning on Day 4, the experimental *S*s were given daily 1-hour habituation sessions, each session consisting of placing *S* in the alley in the *absence* of the object. At the conclusion of each session the object was returned to the alley and the bird was permitted to follow. The control *S*s were treated in the same manner as the experimental *S*s except that beginning on Day 4 the control *S*s were placed in a discriminably different situation for one hour each day. The results indicated that the following exhibited by the experimental *S*s decreased considerably more rapidly than the following exhibited by the control *S*s. There was little doubt that opportunity to habituate to the experimental situation resulted in a marked reduction in the strength of the following response.

There is another implication of the general hypothesis that strength of following is some positive function of the anxiety level present during the period of exposure to the test object. The implication is that an experimental procedure designed to increase anxiety during the test period should increase following. It appears reasonable to assume that electric shock, administered in the absence of the test object but in the presence of stimuli previously associated with following, will produce greater anxiety arousal during subsequent presentation of the object than if the shock had been administered in

the presence of discriminably different stimuli. On this basis we can formulate a second hypothesis: the greater the similarity between cues present during shock and those present during exposure to a test object, the greater the strength at which the following of that object will be maintained.

Evidence relevant to this hypothesis has been presented recently by Moltz, Rosenblum, and Halikas (1959). Peking ducks were again exposed to a cardboard cube for 25 minutes a day for 10 days. Prior to the regularly scheduled exposure of Day 7, the object was removed from the alley and the *S*s were assigned to one of four treatment groups. Each treatment consisted in confining *S* for 15 minutes in a glass compartment the floor of which was constructed of stainless steel rods for the purpose of delivering shock. One half of the *S*s was confined in the compartment when the latter was in the alley; the remaining one half was confined in the compartment when it was outside the alley. The *S*s were further divided into two additional subgroups. One subgroup received a series of brief but intense shocks during the period of its confinement; the other subgroup was not shocked. It was found that *S*s that had been shocked in the alley during confinement were clearly superior to the *S*s of the other subgroups with respect to the strength of following. No significant differences occurred among these other subgroups.

It is also interesting to note that the protocols revealed that superiority in response strength was related to several behavioral indices of anxiety. For example, animals that had been shocked in the alley subsequently emitted a great many distress calls when they were not near the object and would often become startled suddenly for no apparent reason. This startle behavior and the defecation that frequently accompanied it were exhibited considerably less often by the other animals during their exposures to the test object. These results are clearly in accord with theoretical expectation, since they suggest that the anxiety level present during the period of exposure to the test object is an important variable determining the strength at which following will be maintained.

Of course, many other implications of the hypothesis that strength of following is related to anxiety need be examined before the parameters of this relationship become clear. For example, we would want to determine whether strength of following is a linear function of anxiety or whether increasing the level of anxiety would, at some point, result in the disorganization of the response. We would also want to know the extent to which level of anxiety affects the shape of the generalization gradient. In addition, it should be noted that experiments designed to test the hypothesis have employed Peking ducks exclusively, so that a problem arises concerning the generality of the results. It is conceivable that there are differences among precocial avian species with regard to the role of anxiety in maintaining the strength of following. Despite the fact that the following behavior of one species appears to resemble closely that of another (Collias & Collias, 1956), many interspecies comparisons will have to be made before we can conclude that following is influenced by the same variables and serves the same adaptive function in all species in which it occurs. Interesting problems in the phylogeny of behavior would be pre-

sented if it were found that closely related species differed markedly in the extent to which variations in stimulus conditions affected their tendency to maintain the response once it was established.

Why Following Occurs

Lorenz' early statement concerning the uniqueness of what he designated as the imprinting process has been found to be largely unsupported by experimental evidence. We have seen that while the procedure of visually presenting a moving object to a precocial bird of a certain age will almost invariably induce following, there is no empirical justification for assuming that the procedure, as such, directly affects any other response system. We have also seen that the acquisition of the following response is neither rigidly restricted to a specific period in ontogeny nor irreversible once it is established. In addition, we have been able to present experimental evidence which suggests a positive relationship between level of anxiety and the strength at which the following response is maintained. However, we have not as yet discussed what is certainly the most intriguing aspect of the imprinting phenomenon: namely, why following occurs at all.

Lorenz' answer to this question involves a hypothetical releasing mechanism, the initial selectivity of which is assumed to be genetically determined. However, as yet no empirical evidence has been provided which, in the writer's opinion, supports the assumption that such a mechanism is involved in the development of the following response. Indeed, Schneirla maintains that, in general, "claims for innate perceptual schemata have not been validated in adequate experimental investigations" (1957, p. 95). Therefore, disavowing the concept of the releasing mechanism, we turn to examine those processes and events in ontogeny which appear likely to contribute to the development and organization of the following response. In order to provide a framework for the present analysis, let us consider the conditions under which following has been studied in the writer's laboratory and the behavioral events most frequently associated with the occurrence of the response.

The apparatus that has been employed consisted of a well-lighted wooden alley approximately 10-feet long. A motor connected to a pulley system was used to drive the test object about the alley at a constant speed. A duckling between 8 and 10 hours old when placed in this apparatus for the first test trial, usually exhibits very little locomotion. The animal frequently drowses and appears to attend to the object only when it passes near him. During this time none of the conspicuous signs of anxiety is present. However, when the animal is reintroduced to the apparatus 24 hours later for the start of the second 25-minute test trial, it emits many distress calls and frequently runs about the alley for a brief period before beginning to follow. Once following begins, the strength of the response increases progressively until the duckling comes to devote the entire period to pursuing the test object. As long as it remains in close proximity to the object neither distress calls nor startle responses are usually emitted.

What we have just described is the sequence of behavior most often observed to accompany the development of the following response. Variations in this sequence have, of course, also been observed. For ex-

ample, some ducklings do not follow at all, showing a marked avoidance of the object that persists over several trials. Others follow sporadically and never appear to exhibit any signs of emotional arousal when placed in the alley. Some animals will follow during the last few minutes of the first trial and begin to follow immediately at the start of the second trial. Although these ducklings exhibit the sequence of behavior most frequently observed to precede following, the sequence occurs more rapidly. The extent to which these variations are due to differences in physiological and behavioral development at the time of hatching remains to be determined. Perhaps what is most needed in this respect are studies of the growth and activity patterns of the embryo as they relate to the early post-hatching behavior of the neonate. Or, perhaps increased understanding of the variability involved in the onset and termination of the hatching process is what is required. Certainly, a more precise baseline for dating the critical period than that provided by emergence from the egg would, in itself, tend to make for greater homogeneity with respect to the effects of the imprinting procedure. At any rate for the present it seems reasonable to focus attention on the most frequently observed sequence of behavior, deferring any attempt to explain departures from this sequence until such time as further empirical developments permit.

Anxiety and the Development of Following

In addition to the writer, several investigators have called attention to the fact that during a relatively brief period after the occurrence of hatching precocial birds do not exhibit any of the conspicuous signs of emotionality that characterize their later behavior. During this period there is an almost complete absence of such response indices of anxiety as distress calls, avoidance reactions, defecation, etc. It will be recalled that it is precisely at this stage of behavioral development (the critical period) that the bird is exposed initially to a moving object. Now Menner (1938) has pointed out that the fluctuation of retinal illumination that a moving object normally produces in the mammalian eye is magnified in the avian eye by its possession of a highly plicated pecten that projects into the vitreous humor from the point of entrance of the optic nerve. The shadows cast on the retina by the plications of this pecten function to enhance sensitivity to movement by increasing the amount of retinal flicker (i.e., the successive rises and falls in mean illumination of the retina) produced by a moving object. Although, there is no reason, of course, to assume that sensitivity to movement is greater during the critical period than at any other stage in ontogeny, the fact that the bird does possess this sensitivity at a time when its anxiety level is low makes the critical period important with respect to the organization of the following response. Its importance seems to be related entirely to the fact that it provides the occasion for the conjunction or association of a low anxiety drive and an attention-evoking object. If this is indeed the case, then it does not appear unreasonable to suggest that as a consequence of this conjunction the object acquires the capacity to elicit certain autonomically controlled components of the drive state.

A low-anxiety state can be thought of as involving a particular constella-

tion of visceral and cardiac activities, some fraction of which can become conditioned to previously neutral stimuli. It is assumed that during the critical period the (classical) conditioning of these implicit low-anxiety reactions occurs simply by virtue of their association with the object and does not in any way depend upon reinforcement. However, as a result of this conditioning the object acquires the capacity to function as a reinforcer, henceforth mediating new learning. Specifically, it is assumed that when anxiety is subsequently aroused, any response instrumental in bringing the animal into contact with the familiar object will be closely followed in time by anxiety reduction due to the previously acquired capacity of the object to elicit responses incompatible with anxiety. Vigorous pursuit of the object should then continue to occur until either the stimulus situation ceases to arouse anxiety or the reward value of the object is extinguished. The relation between anxiety level and the number of trials over which strong following is maintained has already been discussed.

We have stressed the fact that during the period of initial exposure to the alley and to the test object the bird shows almost no signs of emotional arousal, but that during subsequent exposures distress calls and startle responses become prominent features of its behavior. A problem arises as to why these anxiety responses occur.

It will be recalled that during the period of initial exposure the bird is most often observed to remain in one section of the alley and to drowse at frequent intervals. Relatively few features of the alley, including the moving test object, seem to command the animal's attention. The object appears to dominate the visual environment, while many other stimulus aspects of the apparatus remain "unexplored" and hence unfamiliar. Twenty-four hours later, however, the bird's ability to move about has improved considerably and at the start of the second trial it not only appears more alert but frequently runs about the alley. It does not seem unlikely that the diffuse emotional excitement that the bird exhibits at this time occurs primarily in response to the unfamiliar features of the alley environment with which it now comes into contact.

Recent studies (Hebb, 1946; Jersild, 1954; McBride & Hebb, 1948; Melzack, 1952) of the genesis of emotional behavior in several vertebrate species have shown that strong fear can frequently be elicited by strange, but innocuous, visual stimuli in the absence of any specific avoidance conditioning. What was found necessary for the development of this fear response was a period of early sensory contact with the environment during which time the "familiar" apparently becomes established. Thus, Melzack states that emotional behavior " . . . may appear at any stage of the animal's life when the situation differs greatly from any that the animal has already encountered" (1954, p. 167). While the anxiety responses of the duckling occur very early in ontogeny, there is little doubt that they are frequently evoked by unfamiliar visual stimuli. If the poor locomotor and attentive capacities that the bird displays during the first trial do, in fact, result in limited sensory contact with many aspects of the alley, then it appears reasonable to expect that anxiety will be aroused later when, by virtue of increased activity and general alertness, the bird subsequently comes into contact with

these unfamiliar stimuli. This expectation is supported by results (Hess, 1958) indicating that improvement in locomotor ability is closely accompanied by an increase in the frequency of fear responses during the first 36 hours from the occurrence of hatching.

Without appealing to an innate releasing mechanism that functions exclusively in the organization of the following response, we have been able to suggest an explanation of the manner in which this response comes to be established in ontogeny. It is interesting to note that our conceptualization of the imprinting phenomenon appears to be similar in some respects to the two-factor hypothesis of avoidance learning advanced by Mowrer (1954), Schlosberg (1937), and Solomon and Brush (1956). Both invoke a classical-instrumental conditioning sequence to account for the behavior they seek to explain. Solomon and Brush, for example, conceive of anxiety as being conditioned to a CS by a "process of simple sensory association" (i.e., Pavlovian or classical conditioning) through the contiguity of the CS and the fear reactions aroused by the onset of a noxious stimulus. In addition, they assume that any response subsequently successful in terminating the CS will be followed closely in time by anxiety reduction and thus will be reinforced. We have also focused on the classical conditioning of implicit drive-associated reactions which we have assumed result subsequently in the reinforcement of certain responses through a process of instrumental conditioning. Of course, our concern has been the conditioning of low-anxiety reactions and therefore with the reinforcement of approach rather than with avoidance responses when anxiety is aroused.

Evidence Relevant to the Present Analysis

The value of any theoretical endeavor should be measured not only in terms of its capacity to integrate available data but also in terms of the extent to which it can generate novel (i.e., not previously formulated) functional relationships. It must be determined, therefore, whether the present analysis in terms of anxiety reduction is capable of mediating predictions concerning the influence of certain selected variations in the imprinting procedure on the acquisition and stability of the following response. We turn now to consider some of these predictions.

Characteristics of the Object

It has been noted that the structure of the avian eye increases sensitivity to movement. We have suggested that this sensitivity combined with the limited locomotor and attentive capacities of the newly hatched bird probably causes the moving test object to dominate the visual field—at least during the first trial. However, within the framework of the present analysis, any repetitive visual stimulus that is likely to command the attention of the animal at a time when its anxiety level is low should acquire reward value and subsequently function to reinforce approach responses when anxiety is elicited. Indeed, even a prepotent auditory stimulus should function in this fashion. Evidence in support of the contention that following can be induced by stimuli other than that provided by a moving object has been reported recently by several investigators.

James (1959) placed Plymouth Rock chicks in an alley that contained a flashing light at one end and

a continuous light at the opposite end. The chicks soon developed a strong preference for the intermittent light as evidenced by the greater number of approach responses and by the increasing tendency to remain in the vicinity of the light once it had been approached. Parenthetically, it might be added that James reported the frequency of distress calls decreased markedly when the bird was in the vicinity of the intermittent light. It is also interesting to note that an object consistently associated with the intermittent light subsequently induced following when it moved about the alley. The same object after being paired with the constant light was either ignored or avoided.

With respect to auditory stimuli, Collias and Collias (1956) found that young ducklings of several species will develop a strong tendency to move in the direction of a repetitive low-pitched sound even when the source of the sound is not visible. Fabricius and Boyd (1952–1953) also reported that rhythmic repetition of a brief sound will elicit strong following whether or not it is associated with a moving object. Indeed, some of the ducklings responded more strongly to the retreating sound than to the moving object. However, whether a visual or auditory stimulus was employed, following occurred only after that stimulus had been presented "over a period of varying length."

It has been suggested (Ramsay, 1951) that although a wide variety of inanimate objects can induce following, some are perhaps more effective than others because they bear a closer resemblance to stimuli normally provided by the biological parent. For example, it might be expected that a model decoy of a mature hen would elicit stronger following in a naive bird than a cardboard cube, or that the simulated quacking of an adult duck might serve better than the rhythmic repetition of a nonsense syllable. Such a suggestion implies the existence of a central neural mechanism that is genetically patterned so as to correspond to certain stimulus configurations more closely than to others. It is further implied that the closer the correspondence between extrinsic stimulus and neural structure the more easily the following response can be released.

While we have found no evidence to indicate that innate perceptual schemata of any kind are involved in the development of the following response, the question raised above concerning differential stimulus effectiveness should be examined. Consider the visual modality, and recall that we have already treated the problem of moving as compared with stationary stimuli as it relates to the organization of following. While there has been no attempt to rank order moving objects with respect to the degree to which they resemble the conspecific parent and then to test for effectiveness in inducing following, the limited data that are available indicate that there is no positive relationship. Hinde, Thorpe, and Vince (1956), for example, found that a black box was somewhat more strongly followed than a moorhen model by both moorhens and coots. However, when all the objects employed were considered, they concluded that there was " . . . no indication that any particular characteristic, other than movement, was of especial importance" (p. 224). Fabricius also reports that shape and, within wide limits, size are unrelated either to the occurrence or to the

stability of following. To emphasize the point he exposed ducklings to the " . . . gliding forward movement of a creature with exactly the shape of a swimming female duck" (1951a, p. 276). Not only did this object fail to induce stronger following as compared with other objects, but it was somewhat inferior to the human observer.

In the writer's laboratory, ducklings were exposed either to a green box or to a model decoy of an adult Peking. It was found, contrary to the data reported above, that the decoy elicited stronger following than the box ($p = .08$; two-tailed test). While a subsequent replication failed to confirm this finding, so that it appears to have been the result of sampling error, what theoretical implications would such a difference have entailed had it been confirmed? For example, would it have indicated, to use Ramsay's words, " . . . an innate ability to respond to the biologically correct object"? (1951, p. 12). Perhaps the question should be made even more general: does consistently stronger following (by naive birds) of one object as compared with another, irrespective of whether that object is "biologically correct," necessarily imply the existence of a neural mechanism whose characteristics are in any sense isomorphic with those of the preferred object? We submit that such a preference, if it were to be found, might be more fruitfully analyzed in terms of the excitation of sensory elements in the bird's eye than in terms of the excitation of some specific preformed neural pattern. For example, subsequently greater readiness to follow certain moving objects could be a function of the greater amount of retinal flicker produced by these objects during the critical period by virtue of their size, shape, and rate of movement. It is not unreasonable to expect amount of retinal flicker, or similar receptor-excitation effects, to be related to the extent to which the animal orients toward a particular object and hence to the probability of that object dominating its visual environment. On the basis of the present analysis, we would expect the most salient object during the critical period to have the greatest likelihood of acquiring anxiety-reducing value and of being subsequently followed. In this way differences among objects with respect to their general stimulative effects might become "translated" into preferences that the animal exhibits, making it indeed gratuitous to speak of some preformed "conception" of the preferred object.

Imprinting Effectiveness and the Occurrence of Following

We have conceived the initiation of following as being governed by two processes. The first process consists of the association or conjunction of an attention-evoking stimulus and a low anxiety drive, as a consequence of which certain reaction components of the drive state become conditioned to the stimulus. It is assumed that the conditioning of these low-anxiety reactions causes the stimulus to acquire reward value. The second process involves strengthening responses that serve to bring the animal into contact with this stimulus when anxiety is subsequently aroused.

Now if the two processes that we have described do, in fact, govern the development of following, then it should be possible to manipulate the first process independently of the second. In other words, we would expect the object to acquire reward value independently of the occur-

rence of the following response. Indeed, it should be possible to delay opportunity to follow until well past the critical period, as long as initial exposure occurs during the appropriate time. It is predicted that when following is then permitted, the animal will respond as strongly to the object as it does when it is allowed to govern its own behavior.

There are several studies which appear to support this expectation. For example, Moltz, Rosenblum, and Stettner (unpublished) placed Peking ducks in an apparatus designed specifically to restrict movement while permitting an unobstructed view of an object that traveled about a small area. The *S*s were exposed to the object in this fashion for 25 minutes per day for three days; initial exposure having occurred at approximately 10 hours of age. When subsequently introduced into the alley and allowed to follow this object, they responded in a manner indistinguishable from control *S*s that had been actively pursuing the object for an identical period of time.

James[2] used essentially the same procedure with Plymouth Rock chicks as Moltz et al. had used with ducklings except that a flashing light replaced the moving object. When subsequently permitted to respond, these *S*s approached the light as frequently and as vigorously as control *S*s whose activity had not been previously restricted.

Jaynes (1958a) also presented data that are in accord with the assumption that the acquisition of reward value by an object is not dependent upon the occurrence of overt following. He exposed chicks of various ages to a moving cardboard cube for one 30 minute period. The object was then removed for 10 days after which time *S* was again placed in the apparatus and allowed to follow. Our interest is directed toward those animals that showed no "overt" signs of following during initial exposure. It was found that these *S*s responded somewhat *more* strongly than *S*s that had previously shown some tendency to follow the object. Jaynes used the term "latent imprinting" to emphasize the fact that the effects of exposure to a moving object might not become manifest in behavior until several days later.

A study reported recently by Hess (1957) contains data which do not appear to be in accord with theoretical expectation, since they indicate a relationship between energy expenditure and imprinting effectiveness. Hess exposed mallard ducklings to a model decoy that moved about a circular alley. The distance the decoy traveled, and consequently the energy expended by those ducklings that followed closely, was varied while duration of exposure was held constant. A choice test was then employed involving the familiar model (now stationary) and an unfamiliar model of a different color (either moving or stationary). It was found that increasing the distance over which the duckling was required to follow functioned to increase percentage choice of the familiar model.

Since this finding has been considered to reveal a unique feature of the imprinting process (Thorpe, 1956; Verplanck, 1958) we must carefully examine its relation to the present hypothesis. First, percentage choice was the response measure that Hess employed, while our hypothesis has been formulated primarily in terms of strength of following. The reason for restricting the hypothesis in this way is that percentage choice and

[2] H. James. Personal communication, 1958.

strength of following have been found to provide different estimates of imprinting effectiveness. Jaynes (1958b), for example, reported that chicks which had strongly followed a moving cardboard model subsequently failed to choose consistently between it and a strange object when both were stationary. Rosenblum and the writer have systematically compared the choice responses of ducklings that had strongly followed a moving object with those of ducklings that had either not been permitted to follow during exposure or had not been previously exposed. No significant differences were found.

The apparent failure of the choice technique to yield data that are consistent with strength of following presents an interesting problem. Perhaps the inconsistency is due to the mechanics of the choice technique. Consider the fact that a moving object is suddenly made stationary and is placed in apposition with a strange object. Not only is the test situation altered as a result, but an important stimulus characteristic of the familiar object has been changed; namely, that it is no longer moving. Since such changes are involved in the very nature of the choice technique, it might be expected to provide a somewhat lower estimate of imprinting effectiveness than that provided by a technique that does not require any change in test conditions.

A second point can be raised with respect to Hess' contention concerning the relation between energy expenditure and imprinting effectiveness. On the basis of what has just been said regarding the generalization decrement involved in the use of the choice technique, it would appear reasonable to expect a procedure that allows the object to retain the same speed and pattern of motion throughout the experiment to be superior, or at least equal, to the choice technique in reflecting differences in imprinting effectiveness. Certainly, if Hess' results point to a fundamental feature of imprinting, it should be revealed quite clearly by using strength of following as the response measure. However, when the following response was employed (without involving any change in the stimulus characteristics of the test object) no relation between energy expenditure and imprinting effectiveness was found. As was noted above, birds exposed to, but not permitted to follow, a moving object subsequently followed as vigorously and as strongly as control birds that had been actively pursuing the object during the same period of time. It can, of course, be argued that although the experimental *Ss* were closely confined, some activity (and hence some energy expenditure) was still possible. However, neither Moltz et al. nor James (see Footnote 2) found any relation between the amount of activity that was exhibited and subsequent following. Indeed, the writer observed that some birds sit quietly during confinement and then follow vigorously when the opportunity was afforded. Certainly, these animals expended less energy than their matched controls which at the same time were running about a 10-ft. alley in close pursuit of the object.

The data we have just cited seem to indicate that the effectiveness of imprinting is dependent neither upon energy expenditure in general nor upon the occurrence of overt following in particular. It appears reasonable to maintain, therefore, that Hess' (1958b) statement that "... the strength of imprinting equals the logarithm of the effort expended by the animal during the im-

printing period" (p. 85) requires additional empirical support before it can be generally accepted.

Anxiety Level and the Critical Period

The results of several studies indicated that there is a period in ontogeny during which initial exposure to a moving object is most likely to induce subsequent following. We have maintained that the importance of this period with respect to the organization of the following response is a function, not of some specifically preformed "growth plan," but of the presence of a low-anxiety drive that provides the occasion for the conditioning of certain associated responses. The fact that this drive state is most often prepotent during a relatively brief stage of behavioral development largely limits the effectiveness of initial exposure to the first 15 or 20 hours from the occurrence of hatching. The present analysis suggests, however, that if the sequence of events assumed to be involved in the organization of the following response were to reoccur at a later stage of development, then it should be possible to establish the response in an animal to which the imprinting procedure had not been previously administered. In other words, it is predicted that irrespective of age, the simultaneous occurrence of a low-anxiety drive and a visually dominant (but unfamiliar) object, will result in vigorous following if anxiety is subsequently aroused in the presence of that object either through the introduction of shock or through the introduction of novel (but innocuous) visual stimuli.

There are, of course, many problems involved in testing this hypothesis; the major problem being that of drive manipulation. Recall that a naive duckling older than 20 or 25 hours will often show marked avoidance of a completely strange object, making it difficult to reestablish the initial phase of the assumed drive sequence in the presence of that object. Of course, allowing sufficient time for habituation to occur might result in the desired drive level, the response components of which could then become conditioned to the object. Subsequent exposure, but in an unfamiliar stimulus situation, (which would presumably result in anxiety arousal) might then induce strong following due to the previously acquired anxiety-reducing value of the object. However, if exposure initially elicits fear, an opportunity is provided for components of the fear reaction to become conditioned to the object, this event being incompatible with the drive sequence assumed necessary for the occurrence of following. Gray (1958) has demonstrated that such conditioning can occur quite rapidly and often remains evident for some time. Thus, although data have been obtained (Jaynes, 1957) which indicate that prolonged habituation will frequently result in a "high level of following," a procedure that would eliminate, or at least reduce, the fear manifested initially by juvenile and adult birds in response to a strange object would certainly be preferable. In short, in order to test the hypothesis under consideration we require a method that provides a more "direct" way of reestablishing the low-anxiety state that was present during the critical period.

A promising lead in this respect has been offered recently by Hess (1957) in connection with a study involving the use of such "tranquilizers" as chlopromazine and meprobamate. He reported that these drugs produced a marked reduction of emotionality when administered

to mallard ducklings ranging in age from 24 to 32 hours. Of interest is the fact that the use of chlopromazine served to increase imprinting effectiveness at an age when the imprinting procedure is largely ineffective. It was found that if naive animals, chronologically past the critical period, were exposed to an unfamiliar object while under the influence of chlopromazine, they subsequently (after the drug had worn off) responded more frequently to that object than control animals of the same age to which either water or nembutal had been administered. Unfortunately, while this result is in accord with the present analysis, the fact that meprobamate did *not* prove effective when employed under the same conditions, renders interpretation difficult. Perhaps the contradictory results are due to the widely different pharmacological properties of the two drugs; chlorpromazine being an autonomic suppressant (or a tranquilizer in the narrower sense of the term) and meprobamate a central relaxant (cf. Berger, 1957). Experiments are currently being conducted in the writer's laboratory in the attempt to determine the influence of several different tranquilizing agents on the development of the following response. It is predicted that, irrespective of age, exposure to an unfamiliar test object while under the influence of such agents (especially those that function as "autonomic suppressants") will result in strong following if the animal is subsequently subjected to experimental conditions calculated to arouse anxiety.

The analysis that has been presented in the present paper has been concerned with the manner in which anxiety appears to be related to the development and stability of the following response. Much additional data, obtained through systematic study of a variety of avian species, are required before the parameters of this relationship can be specified precisely. Further experimental work and theoretical effort might also be directed toward determining whether laws governing the acquisition of the following response differ from laws established in relation to the acquisition of those response systems with which contemporary behavior theory is at present concerned (e.g., bar pressing, salivation, etc.). It is believed that the imprinting phenomenon is of sufficient theoretical significance both to learning theorists and to ethologists to warrant such an endeavor.

REFERENCES

ALLEY, R., & BOYD, H. Parent-young recognition in the coot. *Ibis*, 1950, **92**, 46–51.

BEACH, F. A., & JAYNES, J. Effects of early experience upon the behavior of animals. *Psychol. Bull.*, 1954, **51**, 239–263.

BERGER, F. M. The chemistry and mode of action of tranquilizing drugs. *Ann. NY Acad. Sci.*, 1957, **67**, 685–699.

COLLIAS, N. E. The development of social behavior in birds. *Auk*, 1952, **69**, 127–159.

COLLIAS, N. E. The analysis of socialization in sheep and goats. *Ecology*, 1956, **37**, 228–238.

COLLIAS, N. E., & COLLIAS, ELSIE, C. Some mechanisms of family integration in ducks. *Auk*, 1956, **73**, 378–400.

FABRICIUS, E. Ethologie junger Anatiden. *Acta Zool. Fennica*, 1951, **68**, 1–175. (a)

FABRICIUS, E. Some experiments on imprinting phenomena in ducks. *Proc. 10th Int. Orn. Congr.*, 1951, 375–379. (b)

FABRICIUS, E., & BOYD, H. Experiments on the following reactions of ducklings. *Wildfowl Trust Ann. Rep.*, 1952–1953, 84–89.

FANTZ, R. L. Form preferences in newly hatched chicks, *J. comp. physiol. Psychol.*, 1957, **50**, 422–430.

GRAY, P. H. Theory and evidence of imprint-

ing in human infants. *J. Psychol.*, 1958, **46**, 155–166.

HEBB, D. O. On the nature of fear. *Psychol. Rev.*, 1946, **53**, 250–275.

HEDIGER, H. *Wild animals in captivity.* London: Butterworth, 1950.

HEINROTH, O. Beitrage zur Biologie, namentlich Ethologie und Physiologie der Anatiden. *Verhandl. 5th Int. Ornithol. Kongr.*, 1910, 589–702.

HESS, E. H. Effects of meprobamate on imprinting in waterfowl. *Ann. NY Acad. Sci.*, 1957, **67**, 724–732.

HESS, E. H. Evidences for, and theories of, imprinting. Paper presented at American Association for the Advancement of Science, Washington, D. C. December, 1958. (a)

HESS, E. H. "Imprinting" in animals. *Scient. Amer.* March 1958, 81–90. (b)

HINDE, R. A. The following response of moorhens and coots. *Brit. J. anim. Behav.*, 1955, **3**, 121. (Abstract) (a)

HINDE, R. A. The modifiability of instinctive behaviour. *Advanc. Sci., Lond.*, 1955, **12**, 19–24. (b)

HINDE, R. A., THORPE, W. H., & VINCE, M. A. The following response of young coots and moorhens. *Behaviour*, 1956, **9**, 214–242.

JAMES, H. Flicker: An unconditioned stimulus for imprinting. *Canad. J. Psychol.*, 1959, **13**, 59–67.

JAYNES, J. Imprinting: The interaction of learned and innate behavior: I. Development and generalization. *J. comp. physiol. Psychol.*, 1956, **49**, 201–206.

JAYNES, J. Imprinting: The interaction of learned and innate behavior: II. The critical period. *J. comp. physiol. Psychol.* 1957, **50**, 6–10.

JAYNES, J. Imprinting: The interaction of learned and innate behavior: III. Practice effects on performance, retention, and fear. *J. comp. physiol. Psychol.*, 1958, **51**, 234–237. (a)

JAYNES, J. Imprinting: The interaction of learned and innate behavior: IV. Generalization and emergent discrimination. *J. comp. physiol. Psychol.*, 1958, **51**, 238–242. (b)

JERSILD, A. T. *Child psychology.* New Jersey: Prentice Hall, 1954.

LORENZ, K. Der Kumpan in der Umvelt des Vogels. *J. Orn. Lpz.*, 1935, **83**, 137–213; 289–413.

LORENZ, K. The companion in the bird's world. *Auk*, 1937, **54**, 245–273.

LORENZ, K. Morphology and behavior patterns in closely allied species. In B. Schaffner (Ed.), *Group processes.* New York: Josiah Macy Found., 1955. Pp. 168–218.

LORENZ, K. Companionship in bird life. In Claire H. Schiller (Ed.), *Instinctive behavior: The development of a modern concept.* New York: Int. Univer. Press, 1957.

MCBRIDE, A. F., & HEBB, D. O. Behavior of the captive bottle-nose dolphin, *Tursiops truncatus. J. comp. physiol. Psychol.*, 1948, **41**, 111–123.

MELZACK, R. Irrational fears in the dog. *Canad. J. Psychol.*, 1952, **6**, 141–147.

MELZACK, R. The genesis of emotional behavior: An experimental study of the dog. *J. comp. physiol. Psychol.*, 1954, **47**, 166–168.

MELZACK, R., & SCOTT, T. H. The effects of early experience on the response to pain. *J. comp. physiol. Psychol.*, 1957, **50**, 155–161.

MELZACK, R., & THOMPSON, W. R. Effects of early experience on social behaviour. *Canad. J. Psychol.*, 1956, **10**, 82–90.

MENNER, E. Die Bedeutung des Pecten im Auge des Vogels fur die Wahrnehmung von Bewegungen, nebst Bemerkungen uber seine Ontogenie und Histologie. *Zool. Jb., Abt. allg. Zool. Physiol.*, 1938, **48**, 481–583.

MOLTZ, H. Resistance to extinction as a function of variations in stimuli associated with shock. *J. exp. Psychol.*, 1954, **47**, 418–424.

MOLTZ, H., & ROSENBLUM, L. A. Imprinting and associative learning: The stability of the following response in Peking ducks (*Anas platyrhynchous*). *J. comp. physiol. Psychol.*, 1958, **51**, 580–583. (a)

MOLTZ, H., & ROSENBLUM, L. A. The relation between habituation and the stability of the following response. *J. comp. physiol. Psychol.*, 1958, **51**, 658–661. (b)

MOLTZ, H., ROSENBLUM, L. A., & HALIKAS, NINA. Imprinting and level of anxiety. *J. comp. physiol. Psychol.*, 1959, **52**, 240–244.

MOWRER, O. H. Learning theory: Historical review and re-interpretation. *Harvard educ. Rev.*, 1954, **24**, 37–58.

NICE, MARGARET. Some experiences in imprinting ducklings. *Condor*, 1953, **55**, 33–37.

PUMPHREY, R. J. The sense organs of birds. *Ibis*, 1948, **90**, 171–199.

RAMSAY, A. O. Familial recognition in domestic birds. *Auk*, 1951, **68**, 1–16.

RAMSAY, A. O., & HESS, E. H. A laboratory approach to the study of imprinting. *Wilson Bull.*, 1954, **66**, 196–206.

SCHLOSBERG, H. The relationship between success and the laws of conditioning. *Psychol. Rev.*, 1937, **44**, 379–394.

SCHNEIRLA, T. C. Interrelationships of the 'innate' and the 'acquired' in instinctive behavior. In *L'Instinct dans le comporte-*

ment des animaux et de l'homme. Paris: Masson et Cie, 1956. Pp. 387–452.

SCHNEIRLA, T. C. The concept of development in comparative psychology. In Dale B. Harris (Ed.), *The concept of development: An issue in the study of human behavior.* Univer. Minn. Press, 1957. Pp. 78–108.

SPALDING, D. A. Instinct: With original observations on young animals. *MacMillan's Magazine,* 1873, **27**, 282–293. Reprinted in *Brit. J. anim. Behav.*, 1954, **2**, 2–11.

SOLOMON, R. L. & BRUSH, ELINOR. Experimentally derived conceptions of anxiety and aversion. In M. R. Jones (Ed), *Nebraska symposium on motivation.* Univer. Nebr. Press, 1956, Pp. 212–305.

STEVEN, D. M. Transference of "imprinting" in a wild gosling. *Brit. J. anim. Behav.*, 1955, **1**, 14–16.

THORPE, W. H. The evolutionary significance of habitat selection. *J. anim. Ecol.*, 1945, **14**, 67–70.

THORPE, W. H. The learning abilities of birds. *Ibis*, 1951, **93**, 1–52; 252–296.

THORPE, W. H. *Learning and instinct in animals.* Cambridge: Harvard Univer. Press, 1956.

VERPLANCK, W. S. Imprinting and learning. Paper presented at the American Association for the Advancement of Science, Washington, D. C. December, 1958.

(Received April 2, 1959)

Psychological Bulletin
1960, Vol. 57, No. 4, 315-317

CONTRADICTORY CONCLUSIONS FROM TWO SPEED OF PERFORMANCE MEASURES

EUGENE S. EDGINGTON
Kansas State Teachers College, Emporia

Speed of performance can be expressed in terms of amount performed per unit of time or amount of time per unit of performance. The following example will show how different the interpretation of computed statistics can be, for these two alternative measures of speed of performance.

The speeds shown in Table 1 are hypothetical running speeds for two rats, each rat being sampled on two different occasions. The same observations are expressed in two ways in the table. Rat A on the average travelled more feet per second and was, therefore, faster. However, the same rat, Rat A, on the average took more seconds per foot travelled and was, therefore, slower.

TABLE 1

SPEEDS OF RATS EXPRESSED IN TWO WAYS

	Feet per Second		Seconds per Foot (1/Feet-per-Second)	
	Rat A/Rat B		Rat A/Rat B	
	2	4	.50	.25
	8	5	.125	.20
Totals:	10	9	.625	.45
Means:	5	4.5	.3125	.225

One way of settling the contradiction might be to accept the convention of physicists to express velocity only in terms of distance per unit of time, but this solution would not satisfy anyone who considered distance per unit of time and time per unit of distance to be equally valid expressions of speed.

Frequently in research the running times and latency times of rats are subjected to a reciprocal transformation prior to statistical analysis; that is, the statistical analysis is carried out, using the reciprocals of the running times and the reciprocals of the latency times. However, a significant difference between means of the reciprocals of running times cannot be interpreted as indicating a significant difference between means of running times. The same is true of latency times and their reciprocals.

Furthermore, the nonlinearity of the relationship between the two reciprocally related measures ($Y = 1/X$) indicates that there can be inconsistency between them for any statistic that is derived from addition or subtraction of measurement values. Among the statistics subject to inconsistency are these conventional parametric statistics: the mean, the standard deviation, and the product-moment correlation coefficient.

Any computed significance levels that are based on addition or subtraction of the measurement numbers can be different for the two measures; thus, one of the measures might lead to the rejection of a null hypothesis while the other measure might not.

Up to this point the discussion has dealt with running times and latency times of rats, and the reciprocals of these measures. However, the statistical contradictions presented are not the result of carrying out the statistical analysis with the reciprocals instead of the original data. Either of the two reciprocally related measures

is equally valid as a measure of speed of performance, and either one could be considered to be the reciprocal of the other. The point is that there are two precise measures of speed of performance which can lead to contradictory conclusions, for certain statistical computations. This fact applies to speed of performance of any kind. For example, it applies to measuring manual dexterity with a pegboard in terms of the number of pegs per unit of time or in terms of amount of time per peg. Another example is the number of nonsense syllables memorized per unit of time compared to the amount of time per nonsense syllable memorized. And, of course, verbal as well as nonverbal aptitude and achievement tests can employ as a measure of speed either amount of time per unit of performance or amount performed per unit of time. In these and other cases where speed of performance is measured, the significance level obtained can be inconsistent with what would have been obtained had the reciprocally related measure been used.

Proposed Solution

An obvious way to avoid the contradictions that can arise between two reciprocally related measures of speed of performance would be to confine computations to order statistics, e.g., statistics that make use only of the rank order of the original data, not the actual measurement values. The median, mode, and rank correlation coefficient are examples. The objection to this negative kind of solution is that precision of measurement is wasted, and consequently computed statistics would be less sensitive.

There is a solution in which the precision of the measurements can be fully utilized without the possibility of contradiction between the two measures for any computed statistic. A logarithmic transformation of the measurement values (either performance per unit of time or time per unit of performance) prior to statistical analysis accomplishes this result. Since the relationship between the two reciprocally related measures is $Y=1/X$, the relationship between their logarithms is $\log Y=-\log X$; the values are identical, but have opposite signs. Since the numerical values of the logarithms are identical, all statistics computed lead to consistent conclusions. When a statistic is computed for one measure, the conclusion based on it is exactly the same as would be the conclusion based on the same statistic computed for the reciprocally related measure. Any computed level of significance would be identical for any statistic computed for the two measures.

The arithmetic mean of the logarithms is the logarithm of the number that is the geometric mean of the antilogarithms, so a significant difference of means of the logarithms corresponds to a significant difference of geometric means of the antilogarithms. But why should one test for a difference of geometric means when, instead, he is interested in the difference of arithmetic means of the measures? The answer is that he should not test for differences of geometric means instead of arithmetic means if his interest is in the difference of arithmetic means; however, the inconsistency that is associated with the arithmetic mean of speed of performance measurements casts doubt on the meaningfulness and usefulness of hypotheses about the mean speed of performance. In contrast, the geometric mean is consistent for the two measures. A similar argument can be given for the relative meaningfulness of the other statistics that may be computed

from the logarithms of the measurement numbers.

Since most nonparametric statistics would lead to consistent results even without the logarithmic transformation, the principal advantage of the logarithmic transformation is in regard to parametric statistics. It is pertinent, therefore, to consider the normality of the distribution of logarithms of speed of performance measurements. Now, the reciprocal relationship between the two speeds of performance measures requires that for one to be normally distributed the other must not be. This complicates the normality assumption for the speed of performance measures themselves. On the other hand, if the logarithms of one of the two measures are normally distributed, the logarithms of the other measure also must be normally distributed. Another point in favor of the logarithmic transformation is that speed of performance measurements often show a positive skewness, which is consistent with the assumption that their logarithms are normally distributed, since a logarithmic transformation tends to reduce positive skewness.

Psychological Bulletin
1960, Vol. 57, No. 4, 318–328

SIGNIFICANCE TESTS FOR MULTIPLE COMPARISON OF PROPORTIONS, VARIANCES, AND OTHER STATISTICS

THOMAS A. RYAN
Cornell University

Most procedures for multiple comparison which have been published have been concerned only with comparison among the *means* of several samples, and have been further limited by the assumption of normality of the parent populations. A previous paper (Ryan, 1959b) discussed general issues which are involved in all multiple comparisons, whether we are comparing the means of several samples, or we are comparing variances, proportions, medians, correlation coefficients, or any other statistics. These general issues are also independent of any assumptions about the nature of the population distribution or the sampling distribution of the particular statistic we happen to be studying. In that paper, however, concrete procedures for testing multiple comparisons were described only for the case of means. The present paper will describe practical methods for dealing with other statistics, and also for making comparisons by nonparametric methods.

There is a single general method which can be used for all of the problems we have mentioned, merely by "plugging in" the appropriate sampling distribution. We shall call this the "method of adjusted significance levels," and shall illustrate its application to several specific problems. Once the principle of this method is understood, the reader should be able to generalize it to any situation where he knows how to make a two-sample significance test, providing that the available tables of the distribution are complete enough.

In the case of comparing proportions, we can also adapt the methods which have been developed especially for means. Since this provides a slightly more powerful test for proportions than the method of adjusted significance levels, it is to be preferred in this particular case, but it cannot be applied to the other cases. Consequently, we shall begin with the more general procedure.

Method of Adjusted Significance Levels

Like several of the methods for multiple comparison of means, this method is also based upon testing in layers. That is, the highest and lowest sample values are compared first. If they are not found to differ significantly, no further tests are made and we conclude that all samples could have been drawn from the same population. If the highest and lowest sample values are found to differ significantly, we then proceed to test other pairs in order from the outside in. Whenever the extremes of a particular subgroup are found to be nonsignificant, we conclude that there are no significant differences within that subgroup.

Each of the tests is made in exactly the same manner as would be employed for a single pair of samples, but the *nominal* level[1] of significance for the test depends upon the number of samples in the group being compared. Hence the term "adjusted significance levels." By suitable

[1] The *nominal* level of significance refers to the probability value given in the tables which are ordinarily used for comparing two samples alone.

choice of nominal significance levels, we can control the *error rate experimentwise* for the whole set of comparisons. In other words, we are able to state the probability that *one or more* of our conclusions will be incorrect in that we falsely claim significant differences. (See Ryan [1959a] for arguments in support of controlling the error experimentwise).

Specifically, the method of adjusting the significance levels is as follows:

1. If we wish the error rate experimentwise to be a, the first test of the extreme values is made at the nominal level $2a/n(n-1)$, where n is the number of samples being compared. That is, we use whatever tables would ordinarily be used in testing the difference between a single pair of samples, but we enter them for the two-tailed probability of $2a/n(n-1)$ instead of a.

2. If the extremes turn out to be significantly different we test each extreme against the sample next to the other extreme and we do this at the nominal level of $2a/n(n-2)$.

3. When we find a significant difference in Step 2 we proceed to test smaller subgroups. The general rule is that, if we are testing a particular subgroup of k samples, we test the extremes of that subgroup at the nominal level of $2a/n(k-1)$. Notice that this adjusted significance level takes into account both the size of the particular subgroup (k) and the size of the total set of samples (n). (These values of the adjusted significance level apply *only* if the tests are made in layers as described above).

The basis of the method. If the *complete* null hypothesis is true (if all of the samples come from a single population), and if we test all of the differences at the nominal level of $2a/n(n-1)$, the error rate *per experiment*[3] is a. This is because we would be making $n(n-1)/2$ comparisons and the average rate of error per experiment is the product of the number of comparisons times the rate of error per comparison (see Ryan, 1959b, p. 39).

The rate of error per *experiment* is always greater than the error rate *experimentwise*. Thus, if all differences were tested at the nominal rate of $2a/n(n-1)$, the error rate experimentwise would be less than a. In other words, the probability that one or more of the differences will be large enough to be considered significant on this basis will be less than a. Consequently, the probability that the greatest difference will exceed the level required for this nominal rate is less than a.

[3] The rate of error *per experiment* is the expected number of errors per experiment in the long run (Ryan, 1959b). Specifically, if a long series of experiments, each with several comparisons, were actually carried out and we counted the number of erroneous conclusions the three main rates of error would be:

$$\text{rate per comparison} = \frac{\text{no. of } \textit{comparisons} \text{ incorrectly called significant}}{\text{total number of } \textit{comparisons}}$$

$$\text{rate per experiment} = \frac{\text{no. of } \textit{comparisons} \text{ incorrectly called significant}}{\text{total number of } \textit{experiments}}$$

$$\text{rate experimentwise} = \frac{\text{no. of } \textit{experiments} \text{ containing erroneous statements of significance}}{\text{total number of } \textit{experiments}}$$

Thus the rate of error experimentwise can never be greater than the rate per experiment, because we cannot have more experiments with erroneous conclusions than there are errors in individual comparisons.

In the layer method we have described, only the extreme difference is tested at the level $2a/n(n-1)$. The argument of the preceding paragraph shows, however, that this is sufficient to ensure that the rate of error *experimentwise* is less than a when the complete null hypothesis is true. By changing to the experimentwise basis of measuring error rate, we are able to test the internal differences at a less stringent level without overstepping the limit of a for the rate of error. We shall, however, be increasing the rate of error per experiment under the complete null hypothesis unless we make all of the tests at the same nominal level. Working in layers with changing levels of nominal significance as we have proposed increases the power of the test, without sacrificing control of the experimentwise rate of error.

The choice of the nominal level of significance for the tests of the internal layers is based upon various partial null hypotheses. The justification for the level which is to be used can be given for *any* partial null hypothesis as follows: If the complete null hypothesis is not true, then the error rate will depend entirely upon the tests which are made in the internal layers, since it is no longer an error to conclude that the extremes differ significantly. There will be fewer possible errors to be made, and we can therefore be less stringent in testing the internal differences.

Suppose that the true state of affairs consists of several different populations with several samples from each. In general the situation may be represented as:

k_1 samples from Population A
k_2 samples from Population B with $\mu_A > \mu_B$
(μ = population value of the statistic being studied)
k_3 samples from Population C
etc.
$\Sigma k_i = n$ (the total number of samples)

Let us assume that the populations are far enough apart that we can ignore errors which place samples from different populations as significantly different in the wrong order, and that it is almost certain that such samples will be correctly judged significant in the correct order. (This assumption is needed only for clarity of explanation; it will be shown below to be unnecessary in general.)

Following the same argument which we applied to the total group under the complete null hypothesis, the probability of one or more erroneous statements being made about the samples from Population A is:

$$\frac{k_1(k_1-1)}{2}\cdot\frac{2a}{n(k_1-1)}=\frac{k_1}{n}\cdot a$$

The error rate "subgroupwise" would have a corresponding expression for each of the populations. The error rate for the total experiment is not exactly the sum of all of these partial error rates, but fortunately their sum is a good approximation. The complete expression for the error rate experimentwise is:

$$EW \leqq 1-\left(1-\frac{k_1}{n}a\right)\left(1-\frac{k_2}{n}a\right)\left(1-\frac{k_3}{n}a\right)\cdots\left(1-\frac{k_s}{n}a\right)$$

$$=\frac{k_1}{n}a+\frac{k_2}{n}a+\frac{k_3}{n}a\cdots-\frac{k_1k_2}{n^2}a^2-\frac{k_1k_3}{n^2}a^2\cdots+\frac{k_1k_2k_3}{n^3}a^3\cdots$$

If we keep a small, say .05, the terms involving a^2 and higher powers will affect EW very little. In fact, since the squared terms are all negative, EW will be somewhat overestimated if we take only the terms involving the 1st power. Approximately, then, the error rate experimentwise will be:

$$EW \leqq \frac{k_1}{n}a + \frac{k_2}{n}a + \frac{k_3}{n}a + \cdots \frac{k_i}{n}a$$

$$= \frac{\Sigma k}{n}a = a$$

For example, suppose there are 10 samples consisting of 2 samples from each of 5 populations, and take $a = .05$, then $k_1 = 2$, $k_2 = 2$, etc.

$$EW \leqq 1 - \left(1 - \frac{2}{10}.05\right)\left(1 - \frac{2}{10}.05\right)\left(1 - \frac{2}{10}.05\right), \text{ etc.}$$

$$EW \leqq 1 - (1 - .01)^5$$

$$EW \leqq .0490099501$$

Thus the total error rate is divided up among the different groups of samples in proportion to the number of samples from the same population.[8] It would have been better to divide the error rate in proportion to the number of *comparisons* in each group divided by the total number of erroneous *comparisons* which are possible. Since we do not know the latter in the absence of advance knowledge about the true state of affairs, we cannot divide the error rate in this way. It is, however, possible to divide the error rate in proportion to the number of samples involved in each population.

Let us come back now to the problem of erroneous reversal, i.e., where a sample from Population A is called significantly *less* than a sample from Population B, when in fact the conclusion should have been just the reverse. This kind of error is also taken care of by the layer method. Suppose we start with the complete null hypothesis and separate out the errors involving a particular sample value m_A (and its corresponding population value μ_A).

In the first stage of testing the total range of all the samples, there will be three classes of errors which might be made:

1. Conclusions that μ_A is less than the value of μ in some other population.

2. Conclusions that μ_A is greater than some other population value of μ.

3. Conclusions that other pairs of populations are different from each other.

Now suppose that Population A has a higher value of μ than all the rest, while all of the other samples come from a single population. The probability of errors of the first group will then be less than it was under the complete null hypothesis. The conclusions in the second group will no longer be errors, since they state the correct relationship between a pair of populations. Errors in the last group will be somewhat less

[8] For some particular null hypotheses involving a single mean from one of the populations, the general formula for the adjusted significance level is slightly conservative. This is because a portion of the total error rate $(1/n)$ is allotted to this single sample even though there will be no errors involving this mean except for false reversals when the populations are close together. It seemed best to keep the procedure uniform at the expense of this slight loss of power. For example, suppose that five samples come from one population and one from another. The total error rate will be

$$\frac{5.4}{2}\cdot\frac{2a}{6.4} + 0 = \frac{5}{6}a$$

instead of a.

likely than before because they will occur only if some sample from the lower population surpasses the sample from A. As a result the total probability of error is reduced, both because there are fewer possible errors to be made, and because the probability of these errors is reduced.

Confidence limits. The layer method applies only to tests of significance. If we wish to state our results in terms of confidence limits or intervals we pay no attention to the order of the means, the limits being the same whether the means (proportions, etc.) happen to be next to each other or at the extremes of the observed results. The actual sizes of the confidence intervals will vary if the samples differ in size, but otherwise they will all be the same.

As an example suppose that we have five samples with the proportion of occurrence of some event in each. We will make 10 statements of the form: "The difference in proportion between Population A and Population B is between .05 and .42." "The difference in proportion between Population A and Population C is between −.10 and .24," and so on. We can compute these ranges so that the error rate experimentwise is, say, 5%, which means that the probability is .95 that *all* of the statements are correct. If we are dealing with variances of a series of samples, our confidence statements will take the form: "The ratio of variances of Population A and Population B is between 2 and 6." "The ratio of variances of Population A and Population C is between .4 and 1.9," etc. Again these limits can be computed so that the probability is, say, .95 that *all* of the statements are correct.

Confidence limits for multiple comparisons can be made in exactly the same way as for a single comparison of a pair of samples. It is only necessary to adjust the nominal level or error rate per comparison to compensate for the number of confidence statements being made. If the nominal error rate is $2a/n(n-1)$ for each statement, then the error rate for the whole set of statements will be a. This error rate is the rate *per experiment*, and is also an upper limit for the error rate *experimentwise*.

Since the allowance for confidence limits is the same as the difference required in testing the extremes of the group of samples for significance, and there is otherwise no difference in method between multiple and single comparisons, no detailed instructions will be given for obtaining confidence limits. This is not to be taken as implying that significance tests are preferable to confidence limits. In fact, a strong case can be made for the greater usefulness of statements in terms of confidence ranges. We give the methods for significance testing in greater detail because they are more complicated to compute, and also because they are more widely used by psychologists at the present time.

Specific Procedure for Proportions

In order to show how the general method of adjusted significance levels is adapted to various problems, I shall give a "cookbook" outline of the computations required for three different applications, beginning with the problem of proportions.

Steps in computation:

1. The data: n samples consisting respectively of $N_1, N_2 \cdots N_n$ cases. Proportions of occurrence of some event for each sample are $p_1, p_2, \cdots p_n$.

2. Arrange the sample results in order of the size of p, letting p_1 be the smallest sample value and p_n the largest.

TABLE 1

ILLUSTRATION OF MULTIPLE COMPARISONS OF PROPORTIONS AT 5% LEVEL EXPERIMENTWISE

Data

	Samples (Training Methods)					
	B	D	C	A	E	Total
No. of "High Scorers"	12	10	36	36	37	131
No. in sample	60	40	80	60	50	290
Proportion of "High Scorers"	.20	.25	.45	.60	.74	.45

Method of Adjusted Significance Levels

Group	k	Nominal Significance Level $\frac{2\alpha}{n(k-1)}$	z_k	$\bar{p}$	$SE_{dk} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{N_i}+\frac{1}{N_j}\right)}$	$RD_k = z_k \cdot SE_{dk}$	Observed Difference	Significance
B-E	5	.0050	2.81	.45	.0952	.268	.54	s
B-A	4	.0066	2.72	.39	.0890	.242	.40	s
D-E	4	.0066	2.72	.52	.1060	.288	.49	s
B-C	3	.0100	2.58	.32	.0797	.206	.25	s
D-A	3	.0100	2.58	.46	.1017	.262	.35	s
C-E	3	.0100	2.58	.57	.0892	.230	.29	s
BD	2	.0200	2.33	.22	.0846	.197	.05	ns
DC	2	.0200	2.33	.38	.0942	.219	.20	ns
CA	2	.0200	2.33	.51	.0854	.199	.15	ns
AE	2	.0200	2.33	.66	.0906	.211	.14	ns

Tukey Method

Group	k	SR_k	$\overline{SR}_k = \frac{SR_k + SR_n}{2}$	$WSD_k = \overline{SR}_k \cdot SE_{dk}$	RD_k (for (comparison)
B-E	5	2.73	2.73	.260	.268
B-A	4	2.57	2.65	.236	.242
D-E	4	2.57	2.65	.281	.288
B-C	3	2.34	2.54	.202	.206
D-A	3	2.34	2.54	.258	.262
C-E	3	2.34	2.54	.227	.230
BD	2	1.96	2.35	.199	.197
DC	2	1.96	2.35	.221	.219
CA	2	1.96	2.35	.201	.199
AE	2	1.96	2.35	.213	.211

3. Determine the average proportion of occurrence $\bar{p}$ for all samples together by dividing the total number of occurrences by the total number of cases.

4. Using this average value of proportion, $\bar{p}$, find the standard error of the difference between the extreme groups (SE_{dn}) from the usual formula

$$SE_{dn} = \sqrt{\bar{p}(1-\bar{p})(1/N_1 + 1/N_n)}$$

5. Find the standard score z_n in the tables of the normal distribution corresponding to the two-tailed prob-

ability of $2a/n(n-1)$, where a is the experimentwise rate of error desired for the whole set of multiple comparisons.

6. Find the required difference (RD_n) by multiplying the standard score value and the standard error, i.e.:

$$RD_n = SE_{d_n} \cdot z_n$$

7. If the difference between the highest and lowest proportions is less than RD_n we conclude that there is no significant difference among the samples and stop testing. If the difference of the extremes is greater than this required difference, we conclude that the extreme samples differ significantly, and proceed with further tests as follows:

8. Test next the two subgroups obtained by eliminating one of the extreme values, following the procedure of Step 9 for subgroups of size $k = n-1$.

9. Procedure for testing subgroups of k samples:

(*a*) Compute the mean proportion for this particular subgroup.

(*b*) Substitute this for p in the formula for the standard error of the difference, obtaining SE_{dk}.

(*c*) Find z_k from the normal tables corresponding to a two-tailed probability of $2a/n(k-1)$

(*d*) Find $RD_k = SE_{dk} \cdot z_k$; this is the difference required for the extremes of the subgroup to be significant.

10. If either of the subgroups of $n-1$ means is significant by the above procedure we continue by testing those subgroups of $k = n-2$ which are *not* contained in an insignificant group of $n-1$. (Ryan, 1959, b)

The above procedure is illustrated for a specific case of five proportions in Table 1 (middle section).

Application of Tukey's method to proportions. Tukey's procedure[4] for multiple comparison of means is based upon the distribution of the "studentized range," (*SR*), i.e., the range divided by the standard error of the mean. The difference required for significance by Tukey's method is called the *WSD* (wholly significant difference). Assuming that the distribution of sample proportions is approximately normal, we can use the studentized range for multiple comparisons of proportions, making suitable adaptations to allow for the fact that the standard error of the mean is not estimated in the same manner as the standard error of a proportion. Some adjustments are also necessary to allow for samples of unequal numbers.[5]

When these adaptations are carried out, the steps in calculation of *WSD* differ from those described above for *RD* (method of adjusted significance levels) only at certain points. We shall therefore describe only the steps which differ in the two cases, with numbers referring to the outline on previous pages:

1–4. Same as for *RD*.

5. Find the percentage point of the studentized range (SR_n) corresponding to a and to n (the number of samples) and for infinite degrees of freedom. (See Pearson and Hartley, 1954, pp. 176–177).

6. Find *WSD* by multiplying *SR*

[4] J. W. Tukey, The Problem of Multiple Comparisons. Privately circulated monograph, 1953; See also Ryan (1959a, 1959b).

[5] The procedure for means as outlined by Tukey and in my previous paper (Ryan, 1959a) assumed that all samples were of equal size. This is not necessary and the method can be readily modified for use with unequal samples, by working in terms of the standard error of the difference as is done in this outline for multiple comparison of proportions (see Step 6, where the divisor, $\sqrt{2}$, is introduced because of the standard error of the difference).

by the standard error of the difference divided by the square root of 2:

$$WSD_n = \frac{SE_{dn} \cdot SR_n}{\sqrt{2}}$$

7–8. Same, except that WSD_n is used in place of RD_n.

9. (c) We find SR_k (by reading the tables of the studentized range for $n = k$) and average this value with the one previously found for n, i.e.,

$$\frac{SR_k + SR_n}{2} = \overline{SR_k}$$

$$(d)\ WSD_k = \frac{SE_{dk}}{\sqrt{2}} \cdot \overline{SR_k}$$

$$= \frac{SE_{dk}}{\sqrt{2}} \cdot \frac{SR_k + SR_n}{2}$$

10. Same.

The example in Table 1 shows the computations for the *WSD* method in the lower section. It will be noted that there is relatively little difference in the results, but that the Tukey method yields slightly smaller ranges necessary for significance (particularly at the extremes) and is thus slightly more powerful than the method of adjusted significance levels

So far in this discussion of handling sample proportions, we have assumed that the sampling distribution is approximately normal, for both of the methods. In principle, the method of adjusted significance levels could also be used in cases where the samples are too small to justify the normal approximation. In such a case the "Fisher Exact Test" could be applied to the extremes at the adjusted significance level, and so on through the various layers. In practice, however, this is not likely to be very useful because such small samples would not be likely to permit the levels of significance which are required for multiple comparisons.

Multiple Comparison of Variances

If we have computed the variances of a number of different samples, we may wish to draw conclusions about how the corresponding populations differ from one another in variability. The method of adjusted significance levels can also be used for this purpose. The outline of the procedure differs in detail, but not in principle from that which was given above for proportions. In describing the procedure, we shall assume that the F test would be appropriate if only two samples were to be compared.

Steps in computation:

1. Data: n samples of N_1, $N_2 \cdots N_n$ cases respectively, variance estimates from each sample are s_1^2, $s_2^2 \cdots s_n^2$.

2. Arrange the variance in order of magnitude, calling the smallest s_1^2, and the largest s_n^2 (see Table 2 for illustration of the method).

3. Test the extremes s_1^2 and s_n^2 by means of the F test at the nominal level of $2a/n(n-1)$, using the degrees of freedom corresponding to the two sample sizes. If the test is not significant we proceed no further.

N.B.: Tables of the percentage points of the F distribution are usually not extensive enough for this purpose, unless only a small number of samples is involved. In Pearson and Hartley (1954) the percentage points are 25, 10, 5, 2.5, 1, 0.5 and 0.1 *per cent*. Yet, if we have 10 samples to compare and we wish to work at the level of 5% experimentwise the extreme values should be tested at the .1/90 level or .0011 (i.e., 0.11%). This value must be halved in using the percentage points, since the tables are designed for the one-tailed application of F which is appropriate to analysis of variance. The best we could do, then, for 10 variances would be to work at the 10%

TABLE 2

MULTIPLE COMPARISON OF VARIANCES AT 5% LEVEL EXPERIMENTWISE

	Samples				
	B	D	A	C	E
Degrees of Freedom	20	15	25	20	15
Variance (s^2)	10.07	20.35	22.46	39.21	45.63

Summary of Computations

Group	k (no. in layer)	df	Nominal level $= \frac{2\alpha}{n(k-1)}$	Nominal level / 2	Required F	Observed F	Significance
B-E	5	15,20	.00500	.00250	3.95	4.53	s
B-C	4	20,20	.00667	.00333	3.55	3.89	s
D-E	4	15,15	.00667	.00333	4.42	2.24	ns
D-C	3	not tested because D-E not significant					
A-E	3						
B-A	3	25,20	.01000	.00500	3.22	2.23	ns

Note.— No tests possible for $k=2$ since all tests for $k=3$ are insignificant.
Conclusions: There is evidence of only two population variances. B belongs to one, C and E to the other. D and A may belong to either.

level experimentwise. In addition, we would not find the intermediate values which are needed for other layers.

Other percentage points of F can be obtained from the *Tables of the Incomplete Beta Function* (K. Pearson, 1934). The method is stated in Pearson and Hartley (1954, pp. 32–33), and has been used in the computations of the example shown in Table 2.

4. If the extremes are significantly different, proceed to test s_1 vs. s_{n-1} and s_2 vs. s_n, just as we did for proportions, at the nominal significance level of $2\alpha/n(n-2)$.

5. If one or both of the tests in step 2 are significant, continue with other layers until all tests are insignificant for a given size of subgroup. In each subgroup, the test is made at the level of $2\alpha/n(k-1)$. Again it should be noted that, once a particular subgroup is found nonsignificant, no tests are made within that subgroup (Ryan, 1959a).

Note also that the F required for each test depends upon the degrees of freedom in the particular pair being compared. Thus different tests within the same layer will require different F values if the samples are of different size. This, however, is the only adjustment which is needed in applying the method to samples of unequal size.

MULTIPLE NONPARAMETRIC TESTS

Although it is becoming evident that the parametric methods are not so sensitive to nonnormality as was once believed (see Lindquist, 1953, p. 86), there may be situations where we should prefer to use the nonparametric methods. As an example, suppose that we have data which would have been treated by the Mann-Whitney or Festinger method (Festinger, 1946; Mann & Whitney,

TABLE 3

MULTIPLE MANN-WHITNEY TESTS AT 5% LEVEL EXPERIMENTWISE

	Samples Arranged in Order of Medians				
	A	C	B	E	D
Observed Values	7	9	10	11	25
	4	2	7	12	21
	10	10	16	7	14
	7	13	4	14	13
	0	9	23	9	14
	4	11	7	19	7
	8	13	12	16	20
	8	9	13	24	18
Median	7	9.5	12.5	13	16

Summary of Computations

Pair	k	Required level	U (by inspection)	p^a (2-tailed)	Significance
A-D	5	.0050	4	.0020	s
C-D	4	.0067	8	.0100	ns
A-E	4	.0067	5	.0020	s
A-B	3	.0100	14.5	.0740	ns

Note.—Group A is significantly different from Groups D and E. No other differences are significant.

[a] Probabilities read from tables of Mann and Whitney (1947), also given in Siegel (1956). Note that those tables give one-tailed probabilities. For samples larger than 8, probabilities must often be computed from the normal approximation, since percentage points rather than the complete distributions are given.

1955; Siegel, 1956) if we had only two samples, but we happen to have more than two to deal with. A nonparametric "analysis of variance" technique is available (Kruskal & Wallis, 1952; Siegel, 1956, pp. 184–194), but this provides only an overall test and does not give us specific comparisons of samples.

To perform multiple comparisons, we use either the complete distribution of Mann and Whitney's U statistic, or the normal approximation to this distribution. Only a general outline of the procedure will be shown, since it differs little from the previous applications. Table 3 illustrates the computations.

Steps in Computation:

1. Arrange the samples in order of size of the medians.

2. Compute U (the Mann-Whitney statistic based on overlapping of the samples for the two extreme samples.)

3. Test the significance of U at the $2\alpha/n(n-1)$ level, using the exact distribution of U, if available, or the normal approximation of this distribution.[6]

In the latter case the observed value of z is

$$z = \frac{U - \frac{N_1 N_n}{2}}{\sqrt{\frac{N_1 N_n (N_1 + N_n + 1)}{12}}}$$

where N_1 = number of cases in lowest

[6] Available tables (Mann & Whitney, 1955; Siegel, 1956) give the complete distributions up to samples of 8 cases. For larger samples, only the percentage points are given, but the normal approximation is fairly accurate.

sample and N_n = number of cases in highest sample.

4. If the extremes are significantly different continue testing the inner layers at the adjusted significance levels $2a/n(k-1)$ where n is the total number of samples and k is the size of subgroup being tested.

5. Continue in the manner described for proportions and variances.

Summary

A general method for multiple comparisons which is applicable to a variety of measures is presented. This method, called the "method of adjusted significance levels," makes use of the "layer" method of testing and controls the rate of error *experimentwise*. In order to illustrate the manner in which the method is adapted for use with a particular statistic, detailed procedures are outlined for the multiple comparison of proportions, variances, and for one example of nonparametric comparison of sample medians (Mann-Whitney test). Tukey's method for multiple comparison of means which is based upon the distribution of the "studentized range" is also readily adaptable to the multiple comparison of proportions, and instructions for its use are also included in this paper.

REFERENCES

FESTINGER, L. The significance of differences between means without reference to the frequency distribution function. *Psychometrika*, 1946, **11**, 97–105.

KRUSKAL, W. H., & WALLIS, W. A. Use of ranks in one-criterion variance analysis. *J. Amer. Statist. Ass.*, 1952, **47**, 583–621.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin, 1953.

MANN, H. B. & WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. math. Statist.*, 1955, **18**, 50–60.

PEARSON, E. S., & HARTLEY, H. O. *Biometrika tables for statisticians*. Vol. I. Cambridge: Cambridge Univer. Press, 1954.

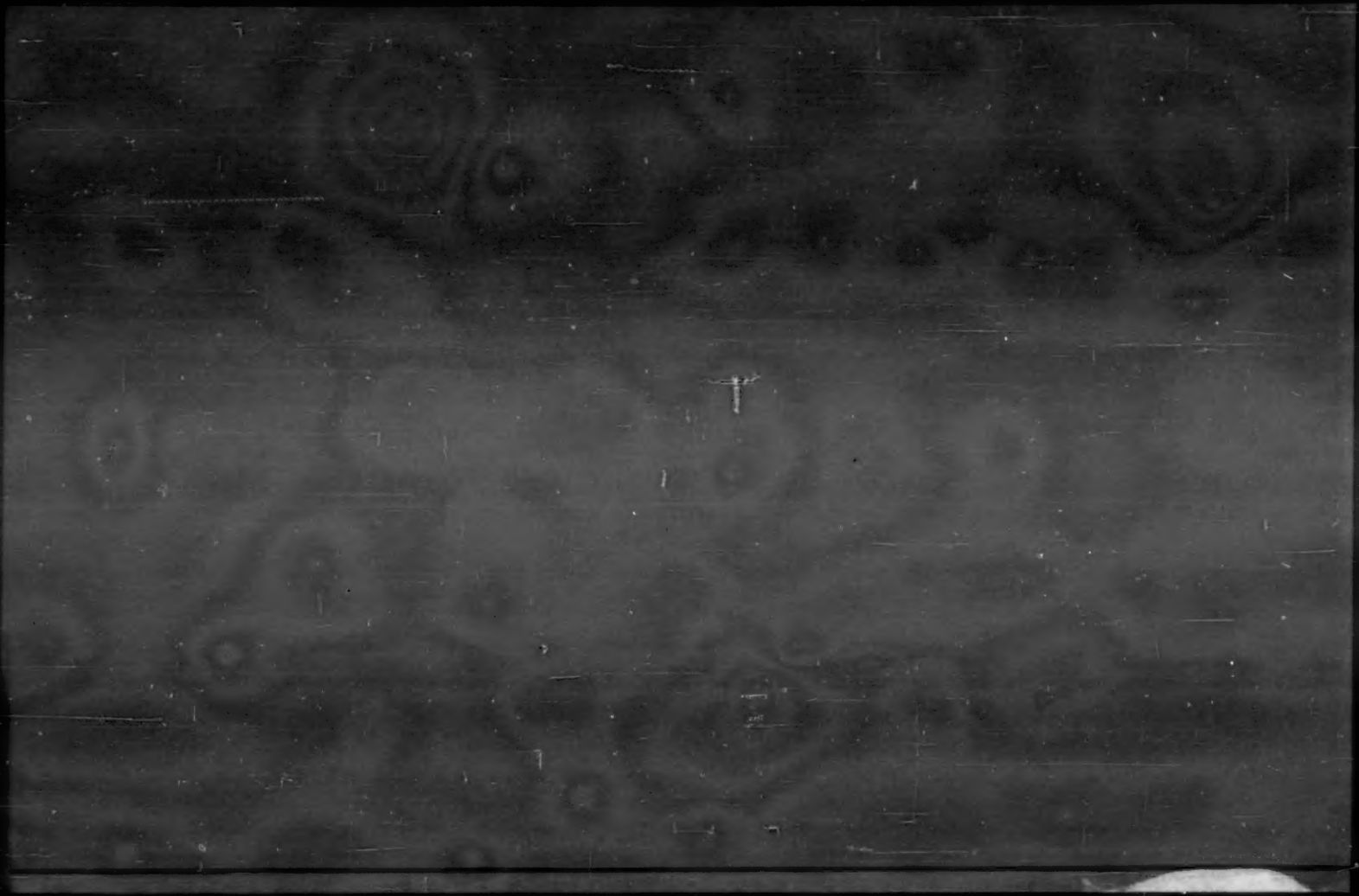
PEARSON, K. Tables of the incomplete Beta-function. London: Cambridge Univer. Press, 1934.

RYAN, T. A. Erratum. *Psychol. Bull.*, 1959, **56**, 396. (a)

RYAN, T. A. Multiple comparisons in psychological research. *Psychol. Bull.*, 1959, **56**, 26–47. (b)

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

(Received June 1, 1959)



PSYCHOLOGICAL MONOGRAPHS

Five new studies—1960, Volume 74

No. 488. **Anoxia, Behavioral Development, and Learning Ability: A Comparative-Experimental Approach**, by G. W. Meier, M. E. Bunch, C. Y. Nolan and C. H. Schaeffer. *Price $1.50*

No. 489. **Similarities in the Structure of Extreme Social Attitudes**, by I. A. Taylor. *Price $1.50*

No. 490. **Perceptual Conditions of Association**, by S. E. Asch, John Ceraso, and Walter Heimer. *Price $1.00*

No. 491. **Meaning and Familiarity in Associative Learning**, by William Epstein, Irvin Rock, and C. B. Zuckerman. *Price $1.00*

No. 492. **Relative Effectiveness of Psychotherapeutic Programs: A Multicriteria Comparison of Four Programs for Three Different Patient Groups**, by G. W. Fairweather et al. *Price $1.00*

Order from:

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth St., N.W.

Washington 6, D.C.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN, U.S.A.